

score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

OM nucleic - nucleic search, using sw model

Run on: October 5, 2004, 14:22:57 ; Search time 11133 Seconds (without alignments)
11216.319 Million cell updates/sec

Title: US-09-970-944-1

Perfect score: 2881

Sequence: 1 agctgggctccgggtgag.....ccttccccacacggggaga 2881

Scoring table: IDENTITY NUC
Gapop 10.0 , Gapext 1.0

Searched: 3470272 seqs, 21671516995 residues

Total number of hits satisfying chosen parameters: 6940544

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database :	GenEmbl.*
1:	gb.ba.*
2:	gb.htg.*
3:	gb.in.*
4:	gb.om.*
5:	gb.ov.*
6:	gb.pat.*
7:	gb.ph.*
8:	gb.pl.*
9:	gb.pr.*
10:	gb.ro.*
11:	gb.sts.*
12:	gb.sy.*
13:	gb.un.*
14:	gb.vi.*
15:	em.ba.*
16:	em.fun.*
17:	em.hum.*
18:	em.in.*
19:	em.mu.*
20:	em.om.*
21:	em.or.*
22:	em.ov.*
23:	em.pat.*
24:	em.ph.*
25:	em.pl.*
26:	em.ro.*
27:	em.sts.*
28:	em.un.*
29:	em.vi.*
30:	em.htg.hum.*
31:	em.htg.inv.*
32:	em.htg.other.*
33:	em.htg.mus.*
34:	em.htg.pln.*
35:	em.htg.rod.*
36:	em.htg.man.*
37:	em.htg.vrt.*
38:	em.sy.*
39:	em.htgo.hum.*
40:	em.htgo.mus.*
41:	em.htgo.other.*

Pred. No. is the number of results predicted by chance to have a

Result No.	Score	Query Match %	Length	DB ID	Description
1	2881	100.0	2881	6	AX527916 Sequence
2	2676.4	92.9	2752	6	AX449572 Sequence
3	2615	90.8	2697	6	AX451652 Sequence
4	2402.6	83.4	3580	6	AX367094 Sequence
5	2327.8	80.8	3992	10	MMU487852
6	2235.8	77.6	3014	6	BD057524 Netrin re
7	2189.4	76.0	2697	6	AX268596 Sequence
8	2189.4	76.0	2697	10	RNU87305
9	2020.6	70.1	3844	10	BC058084 Mus muscu
10	1691.4	58.7	2688	9	BC009333 Homo sapi
11	1570.4	54.5	1787	6	BD057525 Netrin re
12	1291.2	44.8	9700	6	AX054976 Sequence
13	940.2	32.6	9299	10	MMU72634
14	938.4	32.6	2962	5	AY187310
15	936.8	32.5	9328	10	AB118026
16	916	31.8	3646	9	AF055634 Homo sapi
17	895	31.1	3770	9	AY126437 Homo sapi
18	889.4	30.9	4294	10	AK122575
19	889	30.9	2860	6	AX686445
20	887.4	30.8	2860	6	AX686447
21	875.8	30.4	2995	6	AX497288
22	866.4	30.1	2895	6	AX512281
23	859.8	29.8	3672	10	BC057560
24	859.8	29.8	3672	10	BC048162 Mus muscu
25	856.8	29.7	3933	9	AK128132
26	856.8	29.7	3935	9	AB096256 Homo sapi
27	853.6	29.6	3884	6	AX464012
28	853.6	29.6	3884	9	AX358351
29	821.6	28.5	3788	10	MMU487853
30	794.4	27.6	2838	10	RNU87306
31	794.4	27.2	2831	6	BD057526
32	735.8	25.5	813	6	AX054892
33	665.6	23.1	2832	5	AY039459
34	597.6	20.7	2661	6	AX800717
35	597.6	20.7	2868	6	AX800719
36	573.4	19.9	604	6	AX451656
37	560.6	19.5	3252	9	AY358147
38	557.8	19.4	7002	6	BD162193
39	557.8	19.4	7002	9	AB055856
40	553	19.2	3925	10	MMU487854
41	502.6	17.4	1438	6	AX054888
42	467.4	16.2	2406	6	AX833115
43	467.4	16.2	2406	9	AK094595
44	465.8	16.2	1948	6	AX882354
45	465.8	16.2	1948	6	BD159676

ALIGNMENTS

RESULT 1	AX527916	Sequence	1	from Patent WO0229038.	2881 bp	DNA	linear	PAT 21-NOV-2002
LOCUS	AX527916	Sequence	1	from Patent WO0229038.	2881 bp	DNA	linear	PAT 21-NOV-2002
DEFINITION	AX527916	Sequence	1	from Patent WO0229038.	2881 bp	DNA	linear	PAT 21-NOV-2002
ACCESSION	AX527916	Sequence	1	from Patent WO0229038.	2881 bp	DNA	linear	PAT 21-NOV-2002
VERSION	AX527916.1	Sequence	1	from Patent WO0229038.	2881 bp	DNA	linear	PAT 21-NOV-2002
KEYWORDS	AX527916.1	Sequence	1	from Patent WO0229038.	2881 bp	DNA	linear	PAT 21-NOV-2002
SOURCE	AX527916.1	Sequence	1	from Patent WO0229038.	2881 bp	DNA	linear	PAT 21-NOV-2002
ORGANISM	AX527916.1	Sequence	1	from Patent WO0229038.	2881 bp	DNA	linear	PAT 21-NOV-2002
REFERENCE	AX527916.1	Sequence	1	from Patent WO0229038.	2881 bp	DNA	linear	PAT 21-NOV-2002
AUTHORS	AX527916.1	Sequence	1	from Patent WO0229038.	2881 bp	DNA	linear	PAT 21-NOV-2002
TITLE	AX527916.1	Sequence	1	from Patent WO0229038.	2881 bp	DNA	linear	PAT 21-NOV-2002

Pred. No. is the number of results predicted by chance to have a

JOURNAL Patent: WO 0229038-A 1 11-APR-2002;
Curagen Corporation (US)
FEATURES Location/Qualifiers
source 1..2881
/organism="Homo sapiens"
/mol_type="unassigned DNA"
/db_xref="taxon:9606"

ORIGIN
Query Match 100.0%; Score 2881; DB 6; Length 2881;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 2881; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 AGTGGGGCTCCGGGCTGAGCGCTAAAGCGCCCTCCCGCCCGGGGCCCGCGCCG 60
Db 1 AGTGGGGCTCCGGGCTGAGCGCTAAAGCGCCCTCCCGCCCGGGGCCCGCGCCG 60
QY 61 GCCCGCCCGCTGCCCGCCCGCGGCTATGGCGCTCGGCGCGGCTGTGCCAGCGCTCC 120
Db 61 GCCCGCCCGCTGCCCGCCCGCGGCTATGGCGCTCGGCGCGGCTGTGCCAGCGCTCC 120
QY 121 TGGGCATAGTCCTCGCGCTTGCTCCGCGGCTCGGCTGCCAGCTTCCCACTTCTCTGGTGGAGCCGAGG 240
Db 121 TGGGCATAGTCCTCGCGCTTGCTCCGCGGCTCGGCTGCCAGCTTCCCACTTCTCTGGTGGAGCCGAGG 240
QY 181 CMAACCAAGTGCCTGTGGCAACCCGACCTGCTTCCCACTTCTCTGGTGGAGCCGAGG 240
Db 181 CMAACCAAGTGCCTGTGGCAACCCGACCTGCTTCCCACTTCTCTGGTGGAGCCGAGG 240
QY 241 ATGTGTACATCGTCAGAACAGCCAGTGTCTGTGTGCAAGGCGCTGCCCGCACGC 300
Db 241 ATGTGTACATCGTCAGAACAGCCAGTGTCTGTGTGCAAGGCGCTGCCCGCACGC 300
QY 301 AGATCTTCTTCAAGTGCACCGGAGTGTGGTGGCCAGGTGGACACGATGATCGAGCGCA 360
Db 301 AGATCTTCTTCAAGTGCACCGGAGTGTGGTGGCCAGGTGGACACGATGATCGAGCGCA 360
QY 361 GCACAGACGGAGAGTGTGAGCCGACCATGAGGTCCGCACTTAATGTCTCAAGGCAGC 420
Db 361 GCACAGACGGAGAGTGTGAGCCGACCATGAGGTCCGCACTTAATGTCTCAAGGCAGC 420
QY 421 AGGTGAGAGAGTGTTCGGGCTGGAGGAATCTGTGTGCAAGGCGCTGCCCGTCTCT 480
Db 421 AGGTGAGAGAGTGTTCGGGCTGGAGGAATCTGTGTGCAAGGCGCTGCCCGTCTCT 480
QY 481 CGGCGACCAACAGAGTCAAGAGGCTACATCCGATAGCCAGATTCGTGTGCCCTGCCGTC 600
Db 481 CGGCGACCAACAGAGTCAAGAGGCTACATCCGATAGCCAGATTCGTGTGCCCTGCCGTC 600
QY 541 AGCAGAGCGCTGGCCAGGAGGTGTCCCTGGAGCAGGCGATCGTGTGCCCTGCCGTC 600
Db 541 AGCAGAGCGCTGGCCAGGAGGTGTCCCTGGAGCAGGCGATCGTGTGCCCTGCCGTC 600
QY 601 CACCGAGGCGATCCCTCCAGCGAGGTGAGTGTGCTCGGAAACAGAGCCTGTGTGZACC 660
Db 601 CACCGAGGCGATCCCTCCAGCGAGGTGAGTGTGCTCGGAAACAGAGCCTGTGTGZACC 660
QY 661 CGTCCCTGGACCCCAATATATACATCACGCGGAGCACAGCTGTGTGGACAGGCGCC 720
Db 661 CGTCCCTGGACCCCAATATATACATCACGCGGAGCACAGCTGTGTGGACAGGCGCC 720
QY 721 GCCTTGTGACACGGCCAACTACACCTCGTGGCCAGAAACATCGTGGCAGCTGCCCGCA 780
Db 721 GCCTTGTGACACGGCCAACTACACCTCGTGGCCAGAAACATCGTGGCAGCTGCCCGCA 780
QY 781 CGCGCTCGCTGTCTCATCTGTACGTGACGCTGGGTGTGACGTGACCGAGTGTGT 840
Db 781 CGCGCTCGCTGTCTCATCTGTACGTGACGCTGGGTGTGACGTGACCGAGTGTGT 840
QY 841 CCGTCTGAGCGCCAGCTGTGGGCGCGCTGGCAGAAACGAGCGGAGCTGCCAACCC 900
Db 841 CCGTCTGAGCGCCAGCTGTGGGCGCGCTGGCAGAAACGAGCGGAGCTGCCAACCC 900

QY 901 CGGCGCCTCTCAACGGGGGCGCTTCTGTGAGGGGAGAAATGTCCATGACCGACCGTCT 960
Db 901 CGGCGCCTCTCAACGGGGGCGCTTCTGTGAGGGGAGAAATGTCCATGACCGACCGTCT 960
QY 961 CCTCTCTGCTTGTCTCTGTGAGCGAGCTGGAGCCCGTGGAGAAAGTGTGGCGTGTG 1020
Db 961 CCTCTCTGCTTGTCTCTGTGAGCGAGCTGGAGCCCGTGGAGAAAGTGTGGCGTGTG 1020
QY 1021 GGTGTGACTGACACCACTGGCGGAGCGCTGAGTGTCTGTGACCCAGCACCCCGCAACGAG 1080
Db 1021 GGTGTGACTGACACCACTGGCGGAGCGCTGAGTGTCTGTGACCCAGCACCCCGCAACGAG 1080
QY 1081 GGGAGGAGTGCAGGGGCACTGACCTTGACACCCCGCAACTGTACAGTGAACCTCTGTGTAC 1140
Db 1081 GGGAGGAGTGCAGGGGCACTGACCTTGACACCCCGCAACTGTACAGTGAACCTCTGTGTAC 1140
QY 1141 ACAGTGTCTTGGCCCTGAGGAGCTGTGCCCTCTATGTGGGCGCTCATCGCGTGGCGCTCT 1200
Db 1141 ACAGTGTCTTGGCCCTGAGGAGCTGTGCCCTCTATGTGGGCGCTCATCGCGTGGCGCTCT 1200
QY 1201 GCCTGTCTGTGCTGTGTCTCTCATCTCTCTGTTTATTCGCGGAAAGAGGGGCTGG 1260
Db 1201 GCCTGTCTGTGCTGTGTCTCTCATCTCTCTGTTTATTCGCGGAAAGAGGGGCTGG 1260
QY 1261 ACTCAGATGTGGTGTGCTGTCTCACTTCTCACTCAGGCTTCCAGCGCTCAGCATCAAGC 1320
Db 1261 ACTCAGATGTGGTGTGCTGTCTCACTTCTCACTCAGGCTTCCAGCGCTCAGCATCAAGC 1320
QY 1321 CCAGCAAGCAGACAAACCCCATCTGTCTCCCGGAGGATGGGCGCCAGCCCAAGTTCAGACTCACCA 1440
Db 1321 CCAGCAAGCAGACAAACCCCATCTGTCTCCCGGAGGATGGGCGCCAGCCCAAGTTCAGACTCACCA 1440
QY 1381 CCTACAGGGCGTGTCTGTCTCCCGGAGGATGGGCGCCAGCCCAAGTTCAGACTCACCA 1440
Db 1381 CCTACAGGGCGTGTCTGTCTCCCGGAGGATGGGCGCCAGCCCAAGTTCAGACTCACCA 1440
QY 1441 ATGGGCACTGTGACCCCGCTGGTGGCGGCGGCGCACACATGTGCACACAGCTCTCCCA 1500
Db 1441 ATGGGCACTGTGACCCCGCTGGTGGCGGCGGCGCACACATGTGCACACAGCTCTCCCA 1500
QY 1501 CCTCTGAGCGCAGGAGTTCGTCTCCCGCTCTCCACCCAGAACTACTTCCGCTCCCTGC 1560
Db 1501 CCTCTGAGCGCAGGAGTTCGTCTCCCGCTCTCCACCCAGAACTACTTCCGCTCCCTGC 1560
QY 1561 CCGGAGGACACAGCAACATGACCTATGGACCTTCAACTTCTCGGGGCGCGCTGATGA 1620
Db 1561 CCGGAGGACACAGCAACATGACCTATGGACCTTCAACTTCTCGGGGCGCGCTGATGA 1620
QY 1621 TCCCTAATACAGGTATCAGCCTTCTATCCCCCAGATGCCATACCCGAGGAGAGTCT 1680
Db 1621 TCCCTAATACAGGTATCAGCCTTCTATCCCCCAGATGCCATACCCGAGGAGAGTCT 1680
QY 1681 ATGAGATCTACTCACTGACAGCGGAGAGCTGAGGTTCCTTCCCTAGCTGGCTGC 1740
Db 1681 ATGAGATCTACTCACTGACAGCGGAGAGCTGAGGTTCCTTCCCTAGCTGGCTGC 1740
QY 1741 AGACCTCTGAGTCCCATCTGTGAGTGTGAACCCCTGGGCTCTGTCTACCCGCGCAG 1800
Db 1741 AGACCTCTGAGTCCCATCTGTGAGTGTGAACCCCTGGGCTCTGTCTACCCGCGCAG 1800
QY 1801 TCATCTGGCTATGAGCACTGTGGGAGCCAGCCCTGACAGCTGAGCTGGCGCTCA 1860
Db 1801 TCATCTGGCTATGAGCACTGTGGGAGCCAGCCCTGACAGCTGAGCTGGCGCTCA 1860
QY 1861 AAAAGCAGTCTGTGAGGCGAGCTGGGAGCAGGATGTCTCAGCTGGGCGAGGAGGCGC 1920
Db 1861 AAAAGCAGTCTGTGAGGCGAGCTGGGAGCAGGATGTCTCAGCTGGGCGAGGAGGCGC 1920
QY 1921 CCTCCACCTCTACTACTGCGAGCTGGAGGCCAGTGTCTGTCTTCCAGCAGCAGC 1980
Db 1921 CCTCCACCTCTACTACTGCGAGCTGGAGGCCAGTGTCTGTCTTCCAGCAGCAGC 1980
QY 1981 TGGGCGCCTTTGCTGGTGGGAGGCGCTTCAGCTGGGCTGGCGCCAGCGCTCAAGC 2040

[illegible]

Db	1855	CTGCACCTGGCGAGGAGCGCCCTCCCACTCTACTACTGCCAGCTGGAGGCCAGTGCC	1914
Qy	1959	TGCTACGTCCTTTACGGAGCAGCTGGGCGGCTTTGCCCTGTGTGGAGAGGCCCTCAGCGTG	2018
Db	1915	TGCTACGTCCTTTACGGAGCAGCTGGGCGGCTTTGCCCTGTGTGGAGAGGCCCTCAGCGTG	1974
Qy	2019	GCTGCGGCCAAGCGCCTCAAGCTCTTCTGTTTGGCGCGGTGGCGCTGCACCTCCCTCGAG	2078
Db	1975	GCTGCGGCCAAGCGCCTCAAGCTCTTCTGTTTGGCGCGGTGGCGCTGCACCTCCCTCGAG	2034
Qy	2079	TACAACTATCGGGTCTACTGCGCTGCATGACACCCACGATGCACCTCAAGGAGGTGGTGCAG	2138
Db	2035	TACAACTATCGGGTCTACTGCGCTGCATGACACCCACGATGCACCTCAAGGAGGTGGTGCAG	2094
Qy	2139	CTGGAGAAGCAGCTGGGGGACAGCTGATCCAGGAGCCAGGGTCTCTGCACTTCAAGGAC	2198
Db	2095	CTGGAGAAGCAGCTGGGGGACAGCTGATCCAGGAGCCAGGGTCTCTGCACTTCAAGGAC	2154
Qy	2199	AGTTTACCACAACCTGGCGCTATCCATCCAGATGTGCCAGCTCCCTGTGGAAAGAGTAAG	2258
Db	2155	AGTTTACCACAACCTGGCGCTATCCATCCAGATGTGCCAGCTCCCTGTGGAAAGAGTAAG	2214
Qy	2259	CTCCTTGTGAGCTACCAAGAGATCCCTTTTATCACATCTGGAATGGCAGCGACGGTAC	2318
Db	2215	CTCCTTGTGAGCTACCAAGAGATCCCTTTTATCACATCTGGAATGGCAGCGACGGTAC	2274
Qy	2319	TTGCACTGCACCTTACCTTGGAGCGTGTGAGCCCGCAGCATAGTGACCTTGGCGCTGCAAG	2378
Db	2275	TTGCACTGCACCTTACCTTGGAGCGTGTGAGCCCGCAGCATAGTGACCTTGGCGCTGCAAG	2334
Qy	2379	CTGTGGGTGTGGCAGGTGGAGGCGAGCGGCAGAGCTTCAAGCATCACTTCAACATCAACC	2438
Db	2335	CTGTGGGTGTGGCAGGTGGAGGCGAGCGGCAGAGCTTCAAGCATCACTTCAACATCAACC	2394
Qy	2439	AAGGACACAAAGTTTGTGAGCTGCTGGAGTCTGGAGAGTGAAGCGGGGTCCCGAGCCCTG	2498
Db	2395	AAGGACACAAAGTTTGTGAGCTGCTGGAGTCTGGAGAGTGAAGCGGGGTCCCGAGCCCTG	2454
Qy	2499	GTGGGCCCCAGTGCCTTCAAGATCCCTTCTCATTTGGCAGAGAATATAATTTCCAGCCTG	2558
Db	2455	GTGGGCCCCAGTGCCTTCAAGATCCCTTCTCATTTGGCAGAGAATATAATTTCCAGCCTG	2514
Qy	2559	GACCCACCTGTAGGCGGGGTGCCAGTGGCGACCTCTGGCCAGAAACTCCACCTGGAC	2618
Db	2515	GACCCACCTGTAGGCGGGGTGCCAGTGGCGACCTCTGGCCAGAAACTCCACCTGGAC	2574
Qy	2619	AGCCATCTAGCTTTCTTTGCTCCAAAGCCAGCCCGCCAGCCATGATCTCAACCTGTGG	2678
Db	2575	AGCCATCTAGCTTTCTTTGCTCCAAAGCCAGCCCGCCAGCCATGATCTCAACCTGTGG	2634
Qy	2679	GAGCGGGGACATTTCCCAACGGCAACCTCAGCCAGCTGGCTGCAGAGTGGGTGACATG	2738
Db	2635	GAGCGGGGACATTTCCCAACGGCAACCTCAGCCAGCTGGCTGCAGAGTGGGTGACATG	2694
Qy	2739	GGCCAGCCAGCAGCTGGCGCTTTTACAGTGTGGAGGCTCAGTGCTGAGCGCGGCAG	2796
Db	2695	GGCCAGCCAGCAGCTGGCGCTTTTACAGTGTGGAGGCTCAGTGCTGAGCGCGGCAG	2752
RESULT 3			
AX451652			
LOCUS	AX451652	2697 bp	DNA linear PAT 03-JUL-2002
DEFINITION	Sequence 1 from Patent WO0233080.		
ACCESSION	AX451652		
VERSION	AX451652.1		
KEYWORDS	GI:21698597		
SOURCE			
ORGANISM	Homo sapiens (human)		
	Homo sapiens		
	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;		
	Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.		
REFERENCE			
AUTHORS	Koehler, R.H.		
TITLE	Regulation of human netrin binding membrane receptor uncb-1		

1975 TGCACCTCCCTCGAGTACACATCCGGGTCTACTGCTGCATGACACCCACGATGCATC 2034
2124 AAGGAGGTGGTGCAGCTGAGAGAGAGCTGGGGGAGACAGCTGATCCAGAGGACACGGGTC 2183
2035 AAGGAGGTGGTGCAGCTGAGAGAGAGCTGGGGGAGACAGCTGATCCAGGAGGACACGGGTC 2094
2184 CTGCACTTCAAGGACAGTTACACACACCTGCGCTATCCATCCAGATGCGCCAGCTCC 2243
2095 CTGCACTTCAAGGACAGTTACACACACCTGCGCTATCCATCCAGATGCGCCAGCTCC 2154
2244 CTGTGGAAGAGTAAGCTCTGTGACGTACCAAGAGATCCCTTTTATCACATCTGGAAT 2303
2155 CTGTGGAAGAGTAAGCTCTGTGACGTACCAAGAGATCCCTTTTATCACATCTGGAAT 2214
2304 GGCAGCAGCGGTACTTGCATGACACCTTACCTCGAGAGGTGTGAGCCCCAGACACTAGT 2363
2215 GGCAGCAGCGGTACTTGCATGACACCTTACCTCGAGAGGTGTGAGCCCCAGACACTAGT 2274
2364 GACCTGGCCTGCAAGCTGTGGGTGTGGCAGGTGGAGGGCGAGGGGAGAGCTTCAGCATC 2423
2275 GACCTGGCCTGCAAGCTGTGGGTGTGGCAGGTGGAGGGCGAGGGGAGAGCTTCAGCATC 2334
2424 AACTTCAACATCAACCAAGGACACAAAGTTTGTGAGCTCTGCTGAGAGTGAAGCG 2483
2335 AACTTCAACATCAACCAAGGACACAAAGTTTGTGAGCTCTGCTGAGAGTGAAGCG 2394
2484 GGGGTCCAGCCCTGGTGGGCGCCAGTGCCTTCAAGATCCCTTCTCATTTGGCAGAG 2543
2395 GGGGTCCAGCCCTGGTGGGCGCCAGTGCCTTCAAGATCCCTTCTCATTTGGCAGAG 2454
2544 ATAATTTCCAGCTGGAACCCACCTGTAGCGGGGTGCGGACTGCGGAGCTCTGGCCAG 2603
2455 ATAATTTCCAGCTGGAACCCACCTGTAGCGGGGTGCGGACTGCGGAGCTCTGGCCAG 2514
2604 AAATCCACCTGGACAGCCATCTCAGCTTTCTTTGCTTCAAGCCACCCACAGCCATG 2663
2515 AAATCCACCTGGACAGCCATCTCAGCTTTCTTTGCTTCAAGCCACCCACAGCCATG 2574
2664 ATCCCTCAACCTTGGAGCGGGGACCTTCCCGAACGGGAACTCAGCAGCTGGTGCA 2723
2575 ATCCCTCAACCTTGGAGCGGGGACCTTCCCGAACGGGAACTCAGCAGCTGGTGCA 2634
2724 CAGATGGCTGGACTGGGCGAGCCAGACGCTGGCTCTTTCACAGTGTGCGAGGCTGAGTGC 2783
2635 GAGATGGCTGGACTGGGCGAGCCAGACGCTGGCTCTTTCACAGTGTGCGAGGCTGAGTGC 2694
2784 TGA 2786
2695 TGA 2697

RESULT 4
AX367094
LOCUS
DEFINITION
Sequence 13 from Patent WO0198354.
ACCESSION
AX367094
VERSION
AX367094.1 GI:18855296
KEYWORDS
Homo sapiens (human)
SOURCE
Homo sapiens
ORGANISM
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE
1
AUTHORS
Griffin, J.A., Kallik, D.A., Tribouley, C.M., Yue, H., Nguyen, D.B.,
Tang, Y.T., Lal, P., Policky, J.L., Azimzai, Y., Lu, D.A., Graul, R.,
Yao, M.G., Burford, N., Hafalia, A.J., Baughn, M.R., Bandman, O.,
Patterson, C., Yang, J., Xu, Y., Warren, B.A., Ding, L. and
Sanjanwala, M.S.
TITL
Receptors
JOURNAL
Patent: WO 0198354-A 13 27-DEC-2001;
Incyte Genomics, Inc. (US)
FEATURES
Location/Qualifiers

source
1. ...3580
/organism="Homo sapiens"
/mol_type="unassigned DNA"
/db_xref="taxon:9606"
/note="Incyte ID No: 6052371CB1"
ORIGIN
Query Match 83.4%; Score 2402.6; DB 6; Length 3580;
Best Local Similarity 93.4%; Pred. No. 0;
Matches 2615; Conservative 0; Mismatches 9; Indels 177; Gaps 3;
QY 84 GCCATGGCCGCTCGGCGCCGCTGTGGCCAGCGCTCTCTGGGCATAGTCTCGCGCTTGG 143
Db 1 GCCATGGCCGCTCGGCGCCGCTGTGGCCAGCGCTCTCTGGGCATAGTCTCGCGCTTGG 60
QY 144 CTCGCGGGCTCGGGTGCCTCCAGCAGAGTGCACCGTGGCCAAACCCAGTGCCTGGTCCAAC 203
Db 61 CTCGCGGGCTCGGGTGCCTCCAGCAGAGTGCACCGTGGCCAAACCCAGTGCCTGGTCCAAC 120
QY 204 CCGGACCTGCTTCCACCTTCTGTGGAGCCGAGGATGTGTACATCTCAAGTCAAGCAAG 263
Db 121 CCGGACCTGCTTCCACCTTCTGTGGAGCCGAGGATGTGTACATCTCAAGTCAAGCAAG 180
QY 264 CCAGTCTCTGTTGTGTCAAGCCGCTGCGCCACGAGATCTTCTTCAAGTCAAGCAAG 323
Db 181 CCAGTCTCTGTTGTGTCAAGCCGCTGCGCCACGAGATCTTCTTCAAGTCAAGCAAG 240
QY 324 GAGTGGTGGCCAGGTGACACCGTGCATGAGCGGACGACAGCGGGAGCAGTGTGAG 383
Db 241 GAGTGGTGGCCAGGTGACACCGTGCATGAGCGGACGACAGCGGGAGCAGTGTGAG 300
QY 384 CCGACCATGAGGTGCTCGCATTTAATGTCTCAAGCAGCAGGTGTGAGAGGTGTTTCGGGCTG 443
Db 301 CCGACCATGAGGTGCTCGCATTTAATGTCTCAAGCAGCAGGTGTGAGAGGTGTTTCGGGCTG 360
QY 444 GAGGAATACTGTTGCCAGTGCCTGCGCATGAGCTCTCGGGCACCAACAGAGTCAAGAG 503
Db 361 GAGGAATACTGTTGCCAGTGCCTGCGCATGAGCTCTCGGGCACCAACAGAGTCAAGAG 420
QY 504 GCTACATCCGATGACAGATTGGCAAGAACTTCGAGCAGGAGCGCTGGCCACAGGAG 563
Db 421 GCTACATCCGATGACAGATTGGCAAGAACTTCGAGCAGGAGCGCTGGCCACAGGAG 480
QY 564 GTGTCCCTGGACAGGGCATCTGTGCTGCCCTGCGGTCACCGAGGAGCTCCCTCCAGCC 623
Db 481 GTGTCCCTGGACAGGGCATCTGTGCTGCCCTGCGGTCACCGAGGAGCTCCCTCCAGCC 540
QY 624 GAGGTGGAGTGGCTCCGGAAACGAGGACCTGTGGTGGACCCCTGCGACCCCAATGTATAC 683
Db 541 GAGGTGGAGTGGCTCCGGAAACGAGGACCTGTGTGGACCCCTGCGACCCCAATGTATAC 600
QY 684 ATACCGCGGAGACACAGCTGTGTGGACAGGCGCGCTTGTGTACAGCGGCAACTAC 743
Db 601 ATACCGCGGAGACACAGCTGTGTGGACAGGCGCGCTTGTGTACAGCGGCAACTAC 660
QY 744 ACCTGCTGGCCCAAGAAACATCTGTGCACGTCCCGCAGCGCTCCGCTGCTGTCTATCTGTC 803
Db 661 ACCTGCTGGCCCAAGAAACATCTGTGCACGTCCCGCAGCGCTCCGCTGCTGTCTATCTGTC 720
QY 804 TACGTGAACGCTGGGTGGTGTGACGTGGACCGAGTGGTCTGTGACGCGCCAGCTGTGGG 863
Db 721 TAC----- 723
QY 864 CCGCGGTGCGAGAAACGAGGACCGAGTGCACCAACCGCGCGCTTCTCAACGGGCGCT 923
Db 724 ----- 723
QY 924 TTCTGTAGGGGAGAGTGTCCATGATCCAGCGCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 983
Db 724 -----GTGGAC 729
QY 984 GGCAGCTGAGCGCGTGGAGGAGTGGTGGCTGGCTGGCTGGCTGGCTGGCTGGCTGGCTGGCTGG 1043
|||||

Db	730	GGCAGCTGGAGCCCGTGGAGCAAGTGTGGCCCTGTGGCTGGACTGCACCCACTGGCGG	789
Qy	1044	AGCCGTGAGTGTCTGTGACCCAGCACACCCGCCAACGAGGGGAGGAGTGCACGGGCACTGAC	1103
Db	790	AGCCGTGAGTGTCTGTGACCCAGCACACCCGCCAACGAGGGGAGGAGTGCACGGGCACTGAC	849
Qy	1104	CTGGACACCCGCAACTGTACAGTGAACCTCTGTGTACACAGTGTCTTGGCCCTCGAGGAC	1163
Db	850	CTGGACACCCGCAACTGTACAGTGAACCTCTGTGTACACAGTGTCTTGGCCCTCGAGGAC	909
Qy	1164	GTGGCCCTCTATGTGGGCCTCATCGCGTGGCGCCTGTGCCTGGTCTCTGCTGCTGCTTGTCT	1223
Db	910	GTGGCCCTCTATGTGGGCCTCATCGCGTGGCGCCTGTGCCTGGTCTCTGCTGCTGCTTGTCT	969
Qy	1224	CTCATCTCTGTTTATTCGCGAAGAAGGAGGGGTGGACTCAGATGTGGCTGACTCGCTCC	1283
Db	970	CTCATCTCTGTTTATTCGCGAAGAAGGAGGGGTGGACTCAGATGTGGCTGACTCGCTCC	1029
Qy	1284	ATTCTCAGCTCAGCGTTCCAGCCGTCAGCATCAAGGCCACAAAGCAGACAACCCCCAT	1343
Db	1030	ATTCTCAGCTCAGCGTTCCAGCCGTCAGCATCAAGGCCACAAAGCAGACAACCCCCAT	1089
Qy	1344	CTGCTCACCATTCAGCCGCACTCAG--CACCACCAACCTACACGAGGCACTCTCTGT	1400
Db	1090	CTGCTCACCATTCAGCCGCACTCAGCACCAACCAACCACTTACCAAGGCGAGTCTCTGT	1149
Qy	1401	CCCCGGCAGATGGCCCGCAGCCCCAAGTTCAGCTCACCATGGGCACTGCTCAGCCCC	1460
Db	1150	CCCCGGCAGATGGCCCGCAGCCCCAAGTTCAGCTCACCATGGGCACTGCTCAGCCCC	1209
Qy	1461	CTGGTGGCGGCGCACAACCTGCAACAAGTCTTCCCACCTTGAGGCCGAGGAGTTC	1520
Db	1210	CTGGTGGCGGCGCACAACCTGCAACAAGTCTTCCCACCTTGAGGCCGAGGAGTTC	1269
Qy	1521	GTCTCCCGCTCTCCACCGAGACTCTTCGCTCCCTGCCCCGAGGSCACGCAACATG	1580
Db	1270	GTCTCCCGCTCTCCACCGAGACTCTTCGCTCCCTGCCCCGAGGSCACGCAACATG	1329
Qy	1581	ACTTATGGGACTTTCAACTTCTCTCGGGGGCGGTGATGATCCCTAATACAGGTATCAGC	1640
Db	1330	ACTTATGGGACTTTCAACTTCTCTCGGGGGCGGTGATGATCCCTAATACAGGAATCAGC	1389
Qy	1641	CTCCTCATCCCCCAGATGCCATACCCGAGGGAAGATCTATGAGATCTACCTCAGCGTG	1700
Db	1390	CTCCTCATCCCCCAGATGCCATACCCGAGGGAAGATCTATGAGATCTACCTCAGCGTG	1449
Qy	1701	CACAAGCCGGAAGACGTGAGTTGCCCTTACCTAGTGGCTGTGAGACCTGTGAGTCCCATC	1760
Db	1450	CACAAGCCGGAAGACGTGAGTTGCCCTTACCTAGTGGCTGTGAGACCTGTGAGTCCCATC	1509
Qy	1761	GTTAGCTGTGGACCCCTTGGCGCTCTGCTCACCCGGCGAGTCACTCTGGCTATGGACCA	1820
Db	1510	GTTAGCTGTGGACCCCTTGGCGCTCTGCTCACCCGGCGAGTCACTCTGGCTATGGACCA	1569
Qy	1821	TGTGGGAGCCAGCCCTGTACAGTGGAGCCTGCGCCTCAAAAGCAGTCTGTCGAGGGC	1880
Db	1570	TGTGGGAGCCAGCCCTGTACAGTGGAGCCTGCGCCTCAAAAGCAGTCTGTCGAGGGC	1629
Qy	1881	AGCTGGGACGAGATGTGCTGCACTTGGCGAGGAGCGCCCTCCACCTCTACTACTGCTC	1940
Db	1630	AGCTGGG--AGGATGTGCTGCACTTGGGCGAGGAGCGCCCTCCACCTCTACTACTGCTC	1686
Qy	1941	CAGCTGGAGGCGAGTGCCTGTACGCTTTCACCGAGCAGTGGCGCGCTTTCGCTTGGTG	2000
Db	1687	CAGCTGGAGGCGAGTGCCTGTACGCTTTCACCGAGCAGTGGCGCGCTTTCGCTTGGTG	1746
Qy	2001	GGAGAGCCCTCAGCGTGGTGGCGGCAAGCGCCTCAAGCTGCTTCTGTTTGGCGCGGTG	2060
Db	1747	GGAGAGCCCTCAGCGTGGTGGCGGCAAGCGCCTCAAGCTGCTTCTGTTTGGCGCGGTG	1806
Qy	2061	GCCTGCACCTTCCCTCGAGTACACATCCGGGTCTACTGCTTGATGACACCCACGATGCA	2120
Db	1807	GCCTGCACCTTCCCTCGAGTACACATCCGGGTCTACTGCTTGATGACACCCACGATGCA	1866

Qy	2121	CTCAAGGAGGTGGTGCGAGCTGGAGAAAGCAGCTGGGGGACAGCTGATCCAGGAGCCACGG	2180
Db	1867	CTCAAGGAGGTGGTGCGAGCTGGAGAAAGCAGCTGGGGGACAGCTGATCCAGGAGCCACGG	1926
Qy	2181	GTCCTGCACTTTCAGGACAGTTACCAAACTCGCCCTATCCATCCACGATGTGCCAGC	2240
Db	1927	GTCCTGCACTTTCAGGACAGTTACCAAACTCGCCCTATCCATCCACGATGTGCCAGC	1986
Qy	2241	TCCCTGTGGAAGAGTAAGTCTCTTGTGAGCTTACCAGGAGATCCCTCTTTATCACATCTGG	2300
Db	1987	TCCCTGTGGAAGAGTAAGTCTCTTGTGAGCTTACCAGGAGATCCCTCTTTATCACATCTGG	2046
Qy	2301	AATGGCAGCAGCGGTACTTGCACTGCACCTTCACTCCCTGGAGCGTGTGAGCCCCAGCACT	2360
Db	2047	AATGGCAGCAGCGGTACTTGCACTGCACCTTCACTCCCTGGAGCGTGTGAGCCCCAGCACT	2106
Qy	2361	AGTGACCTGGCTCGACAGCTGTGGGTGTGGCAGGTGGAGGGCGACGGGCAGAGCTTCAGC	2420
Db	2107	AGTGACCTGGCTCGACAGCTGTGGGTGTGGCAGGTGGAGGGCGACGGGCAGAGCTTCAGC	2166
Qy	2421	ATCAACTTCAACATCACCAAGGACACAAGGTTTGCTCAGCTGTCTGGCTCTGGAGAGTGAA	2480
Db	2167	ATCAACTTCAACATCACCAAGGACACAAGGTTTGCTCAGCTGTCTGGCTCTGGAGAGTGAA	2226
Qy	2481	CGCGGGTCCAGCCCTGGTGGGCCCCAGTGCCTTCAAGATCCCTTCTCATTCGGCAG	2540
Db	2227	CGCGGGTCCAGCCCTGGTGGGCCCCAGTGCCTTCAAGATCCCTTCTCATTCGGCAG	2286
Qy	2541	AAGATATTTCCAGCCTGGACCCACCTGTGAGCGGGGTGGGACTGGCGGACTCTGGCC	2600
Db	2287	AAGATATTTCCAGCCTGGACCCACCTGTGAGCGGGGTGGGACTGGCGGACTCTGGCC	2346
Qy	2601	CAGAAATCCACCTGGACAGCCATCTCAGCTTCTTTGCTTCCAAAGCCAGCCCAACAGCC	2660
Db	2347	CAGAAATCCACCTGGACAGCCATCTCAGCTTCTTTGCTTCCAAAGCCAGCCCAACAGCC	2406
Qy	2661	ATGATCTCAACCTGTGGAGGCGCGCACTTCCCAAAGGCACTTCAGCCAGCTGGCT	2720
Db	2407	ATGATCTCAACCTGTGGAGGCGCGCACTTCCCAAAGGCACTTCAGCCAGCTGGCT	2466
Qy	2721	GCAGCAGTGGCTGGACTGGGCGACCCAGACGCTGGGCTCTTTCAGTGTGGAGGCTGAG	2780
Db	2467	GCAGCAGTGGCTGGACTGGGCGACCCAGACGCTGGGCTCTTTCAGTGTGGAGGCTGAG	2526
Qy	2781	TGCTGAGCCGCGCCAGCCCGCACCTTACACTCTTCACAGCTTTGGCACCCCAAGGAC	2840
Db	2527	TGCTGAGCCGCGCCAGCCCGCACCTTACACTCTTCACAGCTTTGGCACCCCAAGGAC	2586
Qy	2841	AGGCAGAACCGGCACAGGGGCCCTTCCCCACACCGGGGAGA	2881
Db	2587	AGGCAGAACCGGCACAGGGGCCCTTCCCCACACCGGGGAGA	2627

RESULT 5	
MMU487852	
LOCUS	3992 bp mRNA linear ROD 24-SEP-2002
DEFINITION	Mus musculus mRNA for netrin receptor Unc5hl (Unc5hl gene).
ACCESSION	AJ487852
VERSION	AJ487852.1 GI:22035783
KEYWORDS	netrin receptor Unc5hl; Unc5hl gene.
SOURCE	Mus musculus (house mouse)
ORGANISM	Mus musculus
REFERENCE	1 Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Mus.
AUTHORS	Engelkamp, D.
TITLE	Cloning of three mouse Unc5 genes and their expression patterns at mid-gestation
JOURNAL	Mech. Dev. 118 (1-2), 191-197 (2002)
MEDLINE	22239710
PUBMED	12351186
REFERENCE	2 (bases 1 to 3992)

Db 1702 CTGCCCCGCGTACCAAGCAATGCGCTATGGGAACCTTCAACTCTCTCGGGGCGGCTG 1761
QY 1617 ATGATCCCTTAATACAGGTATCAGCTCTCTATCCCTCCCAAGATGCATACCCGAGGGAAG 1676
Db 1762 ATGATCCCTTAAACAGGAATCAGCTCTCTATACCCCGGAGCCATCCCGGAGGAAG 1821
QY 1677 ATCTATGAGATCTACCTCAGCTGCAACAAGCCGGAAGAGTGGAGTTGCCCTAGCTGGC 1736
Db 1822 ATCTACGAGATCTACTCTACTCTGCAACAGCCAGAGAGTGGAGTTGCCCTAGCTGGC 1881
QY 1737 TGTACAGCCCTGTAGTCCCAATGCTAGTGTGGAACCCCTCTGGGCTCTGTCTACCCGG 1796
Db 1882 TGTACAGCCCTGTAGTCCCAATGCTAGTGTGGAACCCCTCTGGGCTCTGTCTACCCGG 1941
QY 1797 CCAGTCATCTCTGCTATGACCACTGTGGGAGCCAGCCCTGACAGCTGGAGCTGGC 1856
Db 1942 CCAGTCATCTCTGCTATGACCACTGTGGGAGCCAGCCCTGACAGCTGGAGCTGGC 2001
QY 1857 CTCAAAAGCAGTCTGCGAGGCGAGCTGGGAGCAGGATGTCTGCACCTGGCGGAGGAG 1916
Db 2002 CTCAAAAGCAGTCTGCGAGGCGAGCTGGG---AGGAGCTGTGCACCTGGTGGAGG 2058
QY 1917 GCGCCTCCCACTCTACTACTGCGAGCTGGAGGCGAGTCTGTCTCTTCCACCGAG 1976
Db 2059 TCGCCTCTCTCATCTCTACTACTGCGAGCTGGAGGCGGCGCTGTATGTCTTCCACCGAG 2118
QY 1977 CAGCTGGGCGCTTGGCCTGTGGGAGGCGCCCTCAGCGTGGCTGCCAAGCGCCTC 2036
Db 2119 CAGCTAGGCGCTTGGCCTGTGGGAGGCGCCCTCAGCGTGGCTGCCAAGCGCCTC 2178
QY 2037 AAGCTGCTTCTTTGCGCGGTGGCTGCACCTCCCTCGAGTACAAATCCGGGTCTAC 2096
Db 2179 AGGCTCTCTCTTTGCGCCTGTGGCCTGTAGTCCCTCGAGTACAAATCCGGGTCTAC 2238
QY 2097 TGCCTGCATGACACCCACGATGCACTCAAGAGGTGGTGCAGCTGGAGAGCAGCTGGG 2156
Db 2239 TGCCTGCATGACACCCACGATGCACTCAAGAGGTGGTGCAGCTGGAGAGCAGCTGGG 2298
QY 2157 GGACAGCTGATCCAGAGCCAGGCTCTGCACTTCAAGGACAGTACCAGAACCTGGC 2216
Db 2299 GGACAGCTGATCCAGAGCCAGGCTCTGCTGCACTTCAAGGACAGTACCAGAACCTAGT 2358
QY 2217 CTATCCATCCAGATGTGCCAGCTCCCTGTGGAAGAGTAAAGCTCTTGTAGCTACCAAG 2276
Db 2359 CTGTCCATCCAGAGTGGCCAGCTCCCTGTGGAAGAGCAAGCTCTTGTAGCTACCAAG 2418
QY 2277 GAGATCCCTTTTATCAATCTGGAATGSCAGGAGCGGTACTTGCATGCACTTACCC 2336
Db 2419 GAGATCCCTTTTATCAATCTGGAATGSCAGTCACTCAGCAGTATCTGCACTGCACTTACC 2478
QY 2337 CTGGAGCGTGTGAGCCCGCAGACATAGTACCTGGCCTGCAAGCTGTGGGTGGGAGGTG 2396
Db 2479 CTGGAGCGCGTCAATGCGCAGACACGAGCTGGCCTGCAAGGTGTGGGTGGGAGGTG 2538
QY 2397 GAGGCGAGCGGCGAGCTTTCAGATCAACTTCAACATCAACAGGACACAGGTTTGTCT 2456
Db 2539 GAGGAGATGAGCAGAGCTTCAATCACTTAACTTAACTCACTAAGGACACAGGTTTGTCT 2598
QY 2457 GAGCTGTGGCTCTGGAGAGTGAAGCGGGGTCCAGGCCCTGGTGGGCGCCAGTGCCTTC 2516
Db 2599 GAAATGCTGGCTCTGGAGAGTGAAGGGGGGTCCAGGCCCTGGTGGGCGCCAGTGCCTTC 2658
QY 2517 AAGATCCCTCTTCTCATTTCCGCGAGAGATATTTCCAGCTGGACCCCTGTAGGCGG 2576
Db 2659 AAGATCCCTCTTCTCATTTCCGCGAGAGATATTTCCAGCTGGACCCCTGTAGGCGG 2718
QY 2577 GGTGCGGAGCTGTGCGGAGCTCTGCGCCAGAAATCCACCTGGACAGCACTCTCAGCTTCTTT 2636
Db 2719 GCGCGGAGCTGGGAGCTCTAGCCAGAAATCTTCACTTGGACAGCACTTCTTCTTCTTT 2778
QY 2637 GCCTCAAGGCCAGCCCCACAGCCATGATCTCAACTGTGGGAGGCGCGGCACTTCCCC 2696

Db 2779 GCTTCAAGCCCGCCCTACAGCATGATCTTCAACTATGGAGCGCGGCACTTCCCC 2838
QY 2697 AACGGCAACTCAGCCAGCTGGCTGTCAGCAGTGGCTGAGTGGCCAGCCAGCAGCTGGC 2756
Db 2839 AACGGCAACTCAGCCAGCTGGCTGTCAGCAGTGGCTGAGTGGCCAGCCAGCAGCTGGC 2898
QY 2757 CTTCTTCAAGTCTGGAGGCTGAGTGTCTGAGCGCGCCAGGCGCG--ACACTTACTCT 2814
Db 2899 CTTCTTCAAGTCTGGAGGCTGAGTGTCTGAGCGCGCCAGCAGCTGGCTTACTCT 2958
QY 2815 CACCAGCTTTGGCACCACCAAGGACAGGACAGGCGGACAGGCGGCTTCCCCACAC 2874
Db 2959 CACCAGCTTTGGCACCCTGCCAGGACAGCAAAAGCCAGAGGCGGCTTCCCCACAC 3018
QY 2875 GGGGAGA 2881
Db 3019 CCGGGA 3025

RESULT 6
BD057524
LOCUS
DEFINITION
Netrin receptors.
ACCESSION
BD057524
VERSION
BD057524.1 GI:22603130
KEYWORDS
JP 2001505062-A/1.
SOURCE
synthetic construct
ORGANISM
artificial sequences.
REFERENCE
1 (bases 1 to 3014)
AUTHORS
Lavigne, M.T., Leonardo, D.E., Hinck, L., Masu, M. and Masu, K.K.
TITLE
Netrin receptors
JOURNAL
Patent: JP 2001505062-A 1 17-APR-2001;
THE REGENTS OF THE UNIV OF CALIFORNIA
COMMENT
PN JP 2001505062-A/1
PF 17-APR-2001
PF 19-FEB-1998 JP 1998536840
PR 19-FEB-1997 US 08/808982
PI MARC TESSIER LAVIGNE, DAVID E LEONARDO, LINDSAY HINCK, MASAYUKI PI MASU,
PI KAZUKO KEINO MASU
PC C07K14/00, C07K14/00, C07K17/00, C07H21/02, C07H21/04, G01N33/53 CC
Strandedness: Double;
CC Topology: Linear;
FH Key Location/Qualifiers.
source
1..3014
/organism="synthetic construct"
/mol_type="genomic DNA"
/db_xref="taxon:32630"

Query Match 77.6%; Score 2235.8; DB 6; Length 3014;
Best Local Similarity 88.5%; Pred. No. 0;
Matches 2474; Conservative 0; Mismatches 312; Indels 11; Gaps 4;
QY 87 ATGGCGCTCGCGCGCGCTGTGGCCAGCGCTCTCTGGGCAATAGTCTCGCGCTGGCTC 146
Db 1 ATGGCGCTCGCGCGCGCGCTGTGGCCAGTGTCTCTGGGCAATAGTCTCGCGCTGGCTC 60
QY 147 CGCGCTCGGCTGCCAGCAGAGTGCACCGTGGCCACACCCAGTGCCTGGTGCACCCG 206
Db 61 CGTGTTCGGGTGCCAGCAGAGTGCACCGTGGCCAAATCCAGTGCCTGGTGCACCCG 120
QY 207 GACCTGCTTCCCACTTCTCTGGTGGAGCCCGGAGTGTGTACATCGTCAAGAAACAAGCA 266
Db 121 GACCTGCTTCCCACTTCTCTGGTGGAGCCCTGAGAGCGTGTGTACATGTCAAGAAACAAGCG 180
QY 267 GTGCTGCTGTGTGGAAGCGCGTGCCTGCCAGCAGATCTTCTTCAAGTGAACGGGAG 326
Db 181 GTGTTGTTGTGTGCAAGGCTGTGCTGCCACCCAGATCTTCTTCAAGTGAATGGGAA 240
QY 327 TGGGTGCGCAGGTGGACCACTGATCGAGGCGGACGACAGACGGGAGCAGTGTGTAGCCG 386

Db 241 TGGGTCCGCGAGGTGATCATCGTAATTTGAAACGACGACCGACGAGCAGCGGATTTGCCA 300
Qy 387 ACCATGAGGTCCGCAATTAATGTCATCAGGCGAGAGGTGCGAAGGTGTTCCGGCTGGAG 446
Db 301 ACCATGAGGTCCGCAATCAACGATATCGAGGCGAGAGGTAGAGAAAGTGTTCGGCTGGAG 360
Qy 447 GNAATGCTGTCACGTGGTGGCATGAGTCCCTCGGGCACCACCAAGAGTCAGAGGCC 506
Db 361 GAATATGCTGTCACGTGGTGGCATGAGTCCCTCGGGTACCACCAAAAGTCAGAGGCC 420
Qy 507 TACATCCGCATACCGCAGATTGGCGAAGAACTTCGAGCAGAGGCCGCTGGCCAAAGAGGTG 566
Db 421 TACATCCGCATTGGCTATTGTCGCAAGAACTTTGAGCAGAGGCCACTGGCCAAAGAGGTG 480
Qy 567 TCCCTGGAGCAGGGGATCGTGTGCCCTGCGCTGCAACGAGGGGATTCCTCCAGCCGAG 626
Db 481 TCACTGGAGCAAGGCAATTGTACTACCTTGTGCCCCCCCCAGAGGAATCCCCCAGCTGAG 540
Qy 627 GTGGAGTGGCTCCGGAAACGAGGACCTGTGGAGCCGCTCCCTGAGCCCCCAATGTATACATC 686
Db 541 GTGGAGTGGCTTCGAAATGAGGACCTGTGGAGCCCTCCCTCGATCCCAATGTGTACATC 600
Qy 687 AGCGGGGAGCACAGGCTGTGGTGGGACAGAGGCCGCTTGTCTGACACGGCCCAACTACCC 746
Db 601 AGCGGGGAGCACAGGCTAGTCTGTCGTGAGGCCGCTGGCCGACACAGGCCCAACTACCC 660
Qy 747 TCGGTGGCCAGAAACATCGTGGCAAGTCGCGCGAGGCCCTCCGCTGCTCATCGTCTAC 806
Db 661 TGTGTGGCCAGAAACATCGTAGCCGTCGCGAAGCACCTCTGACGCGGTCAATGTTTAT 720
Qy 807 GTGAACGTTGGTGTGCAAGTGGGACGAGTGGTCCGCTGTCAGCGCCAGCTGTGGCGC 866
Db 721 GTGAACGTTGGTGTGCAAGTGGGACGAGTGGTCCGCTGTCAGCGCCAGCTGTGGCGC 780
Qy 867 GGCTGGCAGAAACGGAGCGGAGCTGACCAACCGCGGCCCTCTCAACCGGGCGGCTTTC 926
Db 781 GGCTGGCAGAAACGGAGCGGAGCTGACCAACCGCGGCCCTCTCAACCGGGCGGCTTTC 840
Qy 927 TGTGAGGGCAGAAATGTCATGACCGCACCGTCTCTCTCTGCTGTCTCTGTGGAGCGC 986
Db 841 TGTGAGGGCAGAAATGTCATGACCAACCGCGGCCCTCTCTCTCTGCTGTCTCTGTGGAG 897
Qy 987 AGCTGAGCGCTGGAGCAAGTGTGGCGCTGTGGGCTGAGCTGACACCGCACTGGCGGAGC 1046
Db 898 AGCTGAGTGTGGAGTAAAGTGTGAGCTGTGGGCTTTGATGTCACACCACTGTGGCGGAGC 957
Qy 1047 CGTGAGTGTCTGACCCAGCACCCCGCAACGGAGGGGAGGTGCCAGGGGCACTGACCTG 1106
Db 958 CGCGAGTGTCTGACCCAGCACCCCGCAATGGAGGTGAGGAGTGTGGGGGTGCTGACCTG 1017
Qy 1107 GACACCGCAACTGTACAGTACCTCTGTGTACAGTGTCTTGTGGCCCTGAGGAGCTG 1166
Db 1018 GACACCGCAACTGTACAGTACCTCTGTGTACAGTGTCTTGTGGCCCTGAGGAGCTG 1077
Qy 1167 GCGCTCTATGTGGGCTCATCGCGGTGGCGCTGTGCTGTCTGTCTGTCTGTCTGTCTC 1226
Db 1078 GCTCTCTACATCGGCTTGTGCTGTGGTGTGTGCTCTTCTTGTGTGTGTGGGCTT 1137
Qy 1227 ATCTCTGTTTATTTGCGGAGAGAGGGGCTGGACTCAGATGTGGCTGACTCGTCCATT 1286
Db 1138 GGACTCATTTACTGTGCGAAGAGGAGGGCTGGAATGCGATGTGGCCGACTCGTCCATC 1197
Qy 1287 CTCACCTCAGGCTTCAGCCCTGACGATCAGCCAGCAAGAGCAGACACACCCCATCTG 1346
Db 1198 CTCACCTCAGGCTTCAGGCTGTGACATCAAGCCAGCAAGAGCAGACACACCCACCTG 1257
Qy 1347 CTCACCATCAGCGGACCTCAGCACACCC---ACCACCTTACAGGGGAGTCTCTGTCCC 1403
Db 1258 CTCACCATCAGCGGACCTCAGCACACCCATCACCCTTACAGGGGAGTCTATGTTCG 1317
Qy 1404 CGGCAAGATGGGCGGCGGAGGCTCAAGTTCAGCTCAACATGGGCACTGTCTCAGCCGCTG 1463

Db 1318 AGCAGGATGGACCCAGCCCAAGTTCCAGCTCTCTAATGTGTCACCTGCTCAGCCACTG 1377
Qy 1464 GGTGGCGCGCGCCACACACTGTCACACAGCTCTCCACACTCTGAGGCGAGGAGTTGCTC 1523
Db 1378 GGGAGTGGCGCGCATACGTTGCACACAGCTCACCACACTCTGAGGCTGAGGACTTCGTC 1437
Qy 1524 TCCCGCTCTCTCCACCCAGAACTACTTCCGCTCCCTCGCCCGAGGACACAGCAACATGACC 1583
Db 1438 TCCCGCTCTCTCCACCCAGAACTACTTTCGTTCCCTGCCCCGCGCACAGCAACATGGCC 1497
Qy 1584 TATGGACCTTCAACTCTCTCGGGGCGGCTGATGATCCCTTAATACAGGATACGCTC 1643
Db 1498 TACGGGACCTTCAACTCTCTCGGGGCGGCTGATGATCCCTTAATACGGGATACGCTC 1557
Qy 1644 CTCATCCCCCAGATGCCATACCCGAGGAGATCTATGATCTATACCTCAGCTGAC 1703
Db 1558 CTCATCCCCCAGATGCCATACCCGAGGAAAGATCTAGAGATCTAATCTCAGCTGAC 1617
Qy 1704 AAGCCGGAAGAGCTGAGGTTGCCCTAGCTGGCTCTGACAGCCCTGCTGAGTCCCATGCT 1763
Db 1618 AAGCCGGAAGAGCTGAGGTTGCCCTAGCTGGCTCTGACAGCCCTGCTGAGTCCAGTCT 1677
Qy 1764 AGCTGTGACCCCTGGCGTCTGCTACCCGCGCAGTCTATCTGGCTATGACACCTGCT 1823
Db 1678 AGCTGTGGGCCCCCAGGAGTCTGCTCACCCGCGCAGTCTATCTTGCATGAGCCTG 1737
Qy 1824 GGGAGGCCAGCCCTGACAGCTGGAGCTCGCCCTCAAAAAGCAGCTGTCGAGGGGAGC 1883
Db 1738 GAGAGGCCAGCCCTGACAGCTGGAGTCTGGCCCTCAAAAAGCAGTCTTGGAGGAGT 1797
Qy 1884 TGGGAGCAGGATGTGCTGCACCTGGCGGAGGAGGCGCCCTCCACCTCTACTACTGCCAG 1943
Db 1798 TGGG--AGGATGTGCTGCACCTGGTGAGGAGTCACTTCCACCTCTACTACTGCCAG 1854
Qy 1944 CTGGAGGCCAGTCTGCTGCTACCTCTCACGAGCAGCTGGGCGCTTTCCTGTTGGGA 2003
Db 1855 CTGGAGGCCGCGGCTGCTGCTATGCTTACGAGCAGCTGGGCGCTTTCCTGTTGGGA 1914
Qy 2004 GAGGCCCTCAGGCTGGCTGCGCCAGCGCTCAAGCTCTTCTGTTTTCGCGCGGTGGC 2063
Db 1915 GAGGCCCTCAGGCTGGCTGCCACAGCGCTCAGGCTCTTCTGTTTCTCCGTGGCC 1974
Qy 2064 TGCACCTCCCTCAGGTACAACTCCGGGTCTACTGCTGCATGACACCCAGATGCACTC 2123
Db 1975 TGTACGTCTTGTGTTACAACTCCGAGTGTACTGCTACACGACACCCAGCAGCTCTC 2034
Qy 2124 AAGAGGTGGTGCAGCTGGAGAGCAGCTGGGGGAGCAGCTGATCCAGAGCCACGGTTC 2183
Db 2035 AAGGAGGTGGTGCAGCTGGAGAGCAGCTAGGTGGACAGCTGATCCAGAGGCTCCGCTC 2094
Qy 2184 CTGCACTTCAAGGACAGTTTACCAACCTGCGCTATCCATCCAGATGTGCCAGCTCC 2243
Db 2095 CTGCACTTCAAGGACAGTTTACCAACCTAGTCTCTCATCCACGAGCTGCCAGCTCC 2154
Qy 2244 CTGTGSAAGAGTAAGTCTCTGTGCTACGTACAGGAGATCCCTTTTATCACTTGGAA 2303
Db 2155 CTGTGSAAGAGTAAGTCTCTGTGCTACGTACAGGAGATCCCTTTTATCACATCTGGA 2214
Qy 2304 GGCACGAGCGGTACTTGCAGCTGACCTTCACTGGAGGCTGTCAGCCCGCAGCTAGT 2363
Db 2215 GGCACCGAGGATCTGCACTGSCACTTCACTGGAGGAGTCAACGCCAGCAGCAGC 2274
Qy 2364 GACCTGGCTGCAAGCTGTGGGTGTGGAGGTGGAGGAGGAGGAGGAGGAGGAGGAGG 2423
Db 2275 GACCTGGCTGCAAGGTGTGGGTGTGGAGGTGGAGGAGTGGGAGAGTTCACATC 2334
Qy 2424 AACTTCAACATCAACGAGCAGAGGTTTGTGCTGAGCTGCTGCTCTGAGAGTGAAGCG 2483
Db 2335 AACTTCAACATCACTAAGGACACAGGTTTGTGTAATTTGTGCTCTGAGAGTGAAGGG 2394
Qy 2484 GGGGTCCAGCGCTGTGGGCCCCAGTGTCTCAAGATCCCTTCTCTCTCTCTCTCTCTCT 2543
Db 2395 GGGGTCCAGCGCTGTGGGCCCCAGTGTCTCAAGATCCCTTCTCTCTCTCTCTCTCTCT 2454


```
1464 GGTGGCGCCGCCACACACTGACACACAGCTCTCCCACTCTGAGCGCGAGGAGTTGCTC 1523
1378 GGGAGTGGCGGCATACGTTGACACACAGCTACCCACCTCTGAGGCTGAGGACTTCGTC 1437
1524 TCCCGCTCTCCACACAGAACTACTTCGCTCCCTGCCCGGAGGACACAGCAACATGACC 1583
1438 TCCCGCTCTCCACCAAACTACTTCTGCTCCCTGCCCGGACACAGCAACATGGCC 1497
1584 TATGGAGCTTCAACTTCTCGGGGCGGGCTGATGATCCCTTAATACAGGTATGAGCTTC 1643
1498 TACGGGACCTTCAACTTCTCGGGGCGGGCTGATGATCCCTTAATACAGGGGATCAGCTTC 1557
1644 CTATATCCCCAGATGCCATACCCGAGGAGATCTAGAGATCTACTCTGAGCTGCAC 1703
1558 CTATATCCCCGAGTGCCATCCCCGAGGAGATCTACGAGATCTACTCTACATGAC 1617
1704 AAGCCGGAAGAGCTGAGGTTGCCCCCTAGCTGCTGTACAGCCTGTGAGTCCCATCTGT 1763
1618 AAGCCAGAGAGCTGAGGTTGCCCCCTAGCTGCTGTACAGCCTGTGAGTCCCATCTGT 1677
1764 AGCTGTGACCCCTGGGCTGCTGTACCCGCGCAGTCACTCTGGCTATGACCACTGT 1823
1678 AGCTGTGGCCCCAGGAGTCTGCTGTACCCGCGCAGTCACTCTGCAATGGACCACTGT 1737
1824 GGGAGCCCCAGCTGACAGCTGGAGCTCGCCTCAAAAGCAGTGTGCGAGGCGACG 1883
1738 GGAGAGCCAGCCCTGACAGCTGGAGTCTGCGCCTCAAAAGCAGTGTGCGAGGCGACG 1797
1884 TGGGAGCAGGATGTGCTGACCTGGGCGAGGAGCGCCCTCCCACTCTACTACTGCCAG 1943
1798 TGGG---AGGATGCTGCACCTGTGTGAGGAGTCACTTCCCACTCTACTACTGCCAG 1854
1944 CTGAGGCCAGTGCCTGTGCTGCTTTCACGAGCAGCTGGCGCGCTTTCGCCCTGFGGGA 2003
1855 CTGAGGCCCGGCGCTGTATGTCTTACGAGCAGCTGGCGCGCTTTCGCCCTGFGGGA 1914
2004 GAGGCCCTCAGCTGCTGCGCAGAGCGCTCAAGCTGCTTCTGCTGCGCGGTTGGC 2063
1915 GAGGCCCTCAGCTGCTGCGCAGAGCGCTCAAGCTGCTTCTGCTGCTGCGCGGCT 1974
2064 TGCACCTCCCTCGAGTACAACTCCGGGCTACTGCTGCTGATGACACCCACGATGACATC 2123
1975 TGTAGTCTCTTGTGATACAACTCCGAGTGTACTGCTTACACGACCCACGAGCTCTC 2034
2124 AAGAGGTGGTGTGAGAGAGCTGCTGGGGGACAGCTGATCAGAGAGCCACCGGCT 2183
2035 AAGAGGTGGTGTGAGAGAGCTGATGAGTGTGAGTGTGATCCAGGAGCCTCGCTC 2094
2184 CTGCACTTCAAGGACAGTTACACAACTCGGCTATCCATCCACGATGTGCCAGCTCC 2243
2095 CTGCACTTCAAGGACAGTTACACAACTCGCTTCTCATCCAGAGTGGCCAGCTCC 2154
2244 CTGTGAAGAGTAAGCTCCTGTGCTGCTACAGGAGATCCCTTTTATCAGATCTGGAAT 2303
2155 CTGTGAAGAGTAAGCTCCTGTGCTGCTACAGGAGATCCCTTTTATCAGATCTGGAAT 2214
2304 GGCAGCAGCGGTACTTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 2363
2215 GGCAGCAGCAGTATCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 2274
2364 GACCTGGCTGCAAGCTGTGGGTGTGGCAGTGGAGGCGAGCGGCGAGAGCTTTCAGATC 2423
2275 GACCTGGCTGCAAGCTGTGGGTGTGGCAGTGGAGGAGATGGGAGAGCTTTCAGATC 2334
2424 AACTTCAACATCAACAGGACAAAGTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 2483
2335 AACTTCAACATCAACAGGACAAAGTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 2394
2484 GGGTCCAGCCCTGGTGGGCCCCAGTCTTCAAGATCCCTTCTCTTCTTCTTCTTCTTCT 2543
2395 GGGTCCAGCCCTGGTGGGCCCCAGTCTTCAAGATCCCTTCTCTTCTTCTTCTTCTTCT 2454
```

```
2544 ATAATTCCAGCTGACCCACCTGTAGGCGGGGTGCGGACTGCGGAGTCTTGCCCCAG 2603
2455 ATCATCGCAGTCTGGAACCCCTCTGAGCGGGCGCGGACTGAGAACTCTAGCCCCAG 2514
2604 AAATCCACCTGGACAGCACTCTAGCTTCTTTGCTCTCAAGCCAGCCCAACAGCATG 2663
2515 AAATCCACCTGGACAGCACTCTAGCTTCTTTGCTCTCAAGCCAGCCCAACAGCATG 2574
2664 ATCTCAACCTGTGGAGGCGGGGCACTTCCCAAGCGCAACCTCAGCCAGTCTGCTGCA 2723
2575 ATCTCAACCTGTGGAGGCGGGGCACTTCCCAAGCGCAACCTCAGCCAGTCTGCTGCA 2634
2724 CGAGTGGCTGAGTCTGGGCGGAGCAGCAGCTGTGCTCTTTCACAGTGTGCGAGGCTGAGTGC 2783
2635 GCTGTGGCGGAGTCTGGGCGGAGCAGCAGTGTGCTCTTTCACAGTGTGCGAGGCGGAGTGT 2694
2784 TGA 2786
2695 TGA 2697
```

```
RESULT 8
RN087305
LOCUS
DEFINITION
Rattus norvegicus transmembrane receptor Unc5H1 mRNA, linear ROD 15-MAY-1997
ACCESSION
U87305
VERSION
U87305.1 GI:2055391
KEYWORDS
Rattus norvegicus (Norway rat)
SOURCE
Rattus norvegicus
ORGANISM
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae;
Rattus
REFERENCE
1 (bases 1 to 2697)
Leonardo,E.D., Hinc,L., Masu,M., Keino-Masu,K., Ackerman,S.L. and
Tessier-Lavigne,M.
Vertebrate homologues of C. elegans UNC-5 are candidate netrin
receptors
NATURE 386 (6627), 833-838 (1997)
JOURNAL
MEDLINE
97271897
PUBMED
9126742
REFERENCE
2 (bases 1 to 2697)
Leonardo,E.D., Hinc,L., Masu,M., Keino-Masu,K. and
Tessier-Lavigne,M.
Direct Submission
JOURNAL
SUBMITTED (28-JAN-1997) Anatomy, UCSF, 513 Parnassus, San
Francisco, CA 94143-0452, USA
FEATURES
Location/Qualifiers
1..2697
/organism="Rattus norvegicus"
/mol_type="mRNA"
/db_xref="taxon:10116"
/tissue_type="brain and ventral spinal cord"
/dev_stage="18 day embryo and 13 day embryo"
1..2697
/codon_start=1
/product="transmembrane receptor UNC5H1"
/protein_id="AAB57678.1"
/db_xref="GI:2055392"
```

GVPAIVGPSAFKIPFLIRQIIIASLDPGSRGADWRLTAKLHLDSHLSPFASKPSPT
AMILLNWEARHPNPNGLQLAANAVALGQPDAGLFTVSEAC"

ORIGIN

Query Match 76.0%; Score 2189.4; DB 10; Length 2697;
Best Local Similarity 88.9%; Pred. No. 0;
Matches 2403; Conservative 0; Mismatches 291; Indels 9; Gaps 3;

QY	87	ATGGCCGTCCGGCCCGGCGCTGTGGCCAGAGCGCTCTGGGATATGCTCCGCGCTTGGCTC	146
DB	1	ATGGCCGTCCGGCCCGGCGCTGTGGCCAGAGCGCTCTGGGATATGCTCCGCGCTTGGCTT	60
QY	147	CGCGGCTCGGGTGCACAGAGAGTGCCACCGTGCCCAACCCAGTGCCTGGTGCAACCCG	206
DB	61	CGTGGTTCGGGTGCCACAGAGAGTGCCACCGTGCCCAATCCAGTGCCTGGTGCAACCC	120
QY	207	GACCTGCTTCCCACTTCTCTGGTGAGCCGCGAGGATGTGTACATCCTCAAGAACAGCCA	266
DB	121	GACCTGCTTCCCACTTCTCTGGTGAGCCGCGAGGATGTGTACATCCTCAAGAACAGCCA	180
QY	267	GTGCTGCTTGTGTGCAAGGCGGTGCCGCCACGCGACGCTCTTCAAGTGCAACGGGAG	326
DB	181	GTGTTGTGTTGTGCAAGGCGGTGCCGCCACGCGACGCTCTTCAAGTGCAATGGGAA	240
QY	327	TGGGTGCGCCAGGTGGACCACTGTATGAGCGGACGACAGAGGAGAGTGTGAGCGG	386
DB	241	TGGGTGCGCCAGGTGCATCACGTAAATTGAACGAGCACCGACGAGCGGATTTGCCA	300
QY	387	ACCATGGAGTCCGATTAATGTCTAAGGAGAGGTGAGAGGTTCGGGTGGAG	446
DB	301	ACCATGGAGTCCGATTAACGATATCGAGGCGAGGATGAGAGGATTTGGGTGGAG	360
QY	447	GAATACTGTGTCAGTGGCGTGGCATGGAGCTCTCTGGGCAACCAAGAGTCAGAGGCC	506
DB	361	GAATACTGTGTCAGTGGCATGGAGCTCTCTGGGTAACCAAAAGTCAGAGGCC	420
QY	507	TACATCCGANTAGCAGATTTGGCAAGAACTTTGAGCAGAGCGCGTGGCCAAAGAGTG	566
DB	421	TACATCCGANTAGCCTAATTTGGCAAGAACTTTGAGCAGAGCGCATGGCCAAAGAGTG	480
QY	567	TCCCTGGAGCAGGATCGTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG	626
DB	481	TCACATGGAGCAGGATTTGATACCTTTGTCGCTGCTGCTGCTGCTGCTGCTGCTGCTG	540
QY	627	GTGAGTGGCTCCGGAACGAGGACCTGTGGACCGCTCTGAGCAGAGCGCGTGGCCAAAG	686
DB	541	GTGAGTGGCTTCGAAATGAGGACCTGTGGACCGCTCTGAGCAGAGCGCGTGGCCAAAG	600
QY	687	ACGCGGAGCAGAGCTGTGTGTCGACAGGCGCGCTTGTGACACGCGCAACTACACC	746
DB	601	ACGCGGAGCAGAGCTGTGTGTCGACAGGCGCGCTTGTGACACGCGCAACTACACC	660
QY	747	TGCGTGGCAGAGACATCGTGGCAGCTGCGCGCAGCGCTGCGCTGCTGCTGCTGCTGCTG	806
DB	661	TGTGTGGCAGAAACATCGTAGCCGCTGCGCGCAGCGCTGCTGCTGCTGCTGCTGCTGCTG	720
QY	807	GTGAACCGTGGTGTGTCAGCGTGAACCGAGTGGTCCGTCTGAGCGCGCAGCTGTGGCGC	866
DB	721	GTGAACCGTGGTGTGTCAGCGTGAACCGAGTGGTCCGTCTGAGCGCGCAGCTGTGGCGC	780
QY	867	GGCTGGCAGAAACGAGCGCGAGTGTGCAACCGCGCGCTCTCAACGGGCGCGCTTTC	926
DB	781	GGCTGGCAGAAACGAGCGCGAGTGTGCAACCGCGCGCTCTCAACGGGCGCGCTTTC	840
QY	927	TGTGAGGGCAGATGTTCATGACCGACCGTCTCTCTGCTTGTCTCTGAGCGCGC	986
DB	841	TGTGAGGGCAGATGTTCATGAGCGACCGTCTCTCTGCTTGTCTCTGAGCGCGC	897
QY	987	AGCTGAGCGCGTGTGAGCAAGTGTGGCGCTGTGGCTGAGCTGACCGACCTGCGGAGC	1046
DB	898	AGCTGAGCTTGTGAGGTAAGTGTGAGCGCTGTGGCTTGTGAGCTGACCGACCTGCGGAGC	957
QY	1047	CGTGAAGTCTCTGACCCAGCACCCCGCAACGAGGGGAGGAGTGCAGGGCACTGACCTG	1106

DB	958	CGGAGTGTCTGACCCAGACCCCGCAATGGAGGTGAGAGTGTCCGGGTGTGACCTG	1017
QY	1107	GACACCGCAACTGTACCACTGACCTCTGTGTACACAGTGTCTTCTGCGCCCTGAGGACGTG	1166
DB	1018	GACACCGCAACTGTACCACTGACCTCTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT	1077
QY	1167	GCCTCTATGTGGGCTCATATGCGCGTGGCGCTGTGCTGGTCTGCTGCTGCTGCTGCTGCTGCTG	1226
DB	1078	GCTCTCTATACATGGGCTTGT	1137
QY	1227	ATCTCTGTTTATTTGCGGAGAGAGGAGGAGGCTGGAATCAGATGTGCTGACTGCTGCTGCTGCTG	1286
DB	1138	GGACTCATTTACTGTGCGAAGAGGAGGAGGCTGGAATCAGATGTGCTGACTGCTGCTGCTGCTG	1197
QY	1287	CTCACCTCAGGCTTCCAGGCGCTCAGCATCAAGCCCGCAGCAAAAGAGCAGACAAACCCCATCTG	1346
DB	1198	CTCACCTCAGGCTTCCAGGCTTCCAGCATCAAGCCCGCAGCAAAAGAGCAGACAAACCCCATCTG	1257
QY	1347	CTCACCTCAGGCGGACCTCAGCAGCACCC--ACCACCTACAGGGCAGTCTCTGTGCC	1403
DB	1258	CTCACCTCAGGCGGACCTCAGCAGCACCCACCTACACCTACAGGGCAGTCTATGTTCC	1317
QY	1404	CGGAGTGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG	1463
DB	1318	AGGAGGATGGAGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG	1377
QY	1464	GGTGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG	1523
DB	1378	GGGAGTGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG	1437
QY	1524	TCCGCGCTCTCACCGGAGACTTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG	1583
DB	1438	TCCGCGCTCTCACCGGAGACTTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG	1497
QY	1584	TATGGAGCTTCAACTTCTCTGGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG	1643
DB	1498	TACGGGAGCTTCAACTTCTCTGGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG	1557
QY	1644	CTCATCCCCCAGATGSCATATACCCGAGGAGAGTCTATGAGATCTACTACGCTGAC	1703
DB	1558	CTCATCCCCCAGATGSCATATACCCGAGGAGAGTCTATGAGATCTACTACGCTGAC	1617
QY	1704	AAGCGGAGAGCTGAGGTGCGCTTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG	1763
DB	1618	AAGCGGAGAGCTGAGGTGCGCTTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG	1677
QY	1764	AGCTGTGACCCCTGCGCTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG	1823
DB	1678	AGCTGTGACCCCTGCGCTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG	1737
QY	1824	GGGAGCCGACCTTGACAGCTGGAGCTTGGCGCTCAAAAAGCAGTCTGCTGCGAGGCGAGC	1883
DB	1738	GGGAGCCGACCTTGACAGCTGGAGCTTGGCGCTCAAAAAGCAGTCTGCTGCGAGGCGAGT	1797
QY	1884	TGGGAGCAGGATGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG	1943
DB	1798	TGGG---AGGATGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG	1854
QY	1944	CTGAGGCGCAGTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG	2003
DB	1855	CTGAGGCGCAGTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG	1914
QY	2004	GAGGCGCTTACGCTGGCTGCGCGCAAGCGCTCAAGCTGCTTCTGCTTGTGCGCGGCTGCGC	2063
DB	1915	GAGGCGCTTACGCTGGCTGCGCGCAAGCGCTCAAGCTGCTTCTGCTTGTGCTGCTGCTGCTGCTG	1974
QY	2064	TGCACTCTCTGAGTACAACTGCGGCTTACTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG	2123
DB	1975	TGTACGCTCTTGTGATCAACATCCGAGTGTACTGCTTACACGACCCCAACGACGCTCTC	2034
QY	2124	AAGGAGTGTGAGCTGGAGAGCAGTGGGGGAGAGTGTGAGGAGGAGTGTGAGGAGGAGTGTGAGGAGG	2183

GTQYLHCTFLTLERNASTSDLACKVWVQVGDGOSFNINFTKDTFAEMIALES
EGVPAIVGSPAFKIPILIRQIITSLDPCSRGADWRTLAOKHLDSLSPSPKPS
PTAMINLWEARHPNGLNLQAAAAVAGLQDPDAGLFTVSEAC"
714. .926
/note="IG; Region: Immunoglobulin"
/db_xref="CDD:smart00409"
misc_feature 984. .1130
/note="TSP1; Region: Thrombospondin type 1 repeats"
/db_xref="CDD:smart00209"
misc_feature 1566. .1877
/note="ZU5; Region: Domain present in ZO-1 and Unc5-like
necrin receptors"
/db_xref="CDD:smart00218"
misc_feature 2496. .2684
/note="DEATH; Region: DEATH domain, found in proteins
involved in cell death (apoptosis). Alpha-helical domain
present in a variety of proteins with apoptotic functions.
Some (but not all) of these domains form homotypic and
heterotypic dimers"
/db_xref="CDD:smart00005"
ORIGIN

Query Match 70.1%; Score 2020.6; DB 10; Length 3844;
Best Local Similarity 84.1%; Pred. No. 0;
Matches 2429; Conservative 0; Mismatches 274; Indels 184; Gaps 6;

QY 1 AGCTGGGGCTCGGGCTGAGGGCT-AAAGCCGCTCCCGCGCGGGGGCCCCGGGCC 59
Db 169 AGGTGGGGCTCTCGGCTGAGGGCTGACAGCTCTCTCCCGCGCGGGGGCCCCTAGTCC 228
QY 60 GGCCTGGCGCTGCGCGCGCGCCATGGCCGTCGCGCGCGGGCTGTGGCCAGCGTCC 119
Db 229 AGCCCGCTCGT-----CCGCGCGCGCGCCATGGCCGTCGCGCGCGGGCTGTGGCCAGCGTCC 284
QY 120 CTGGSCATAGTCTCGCGCTTGGTTCGCGCGCTCGGGTGCCAGAGTGCACCGT 179
Db 285 CTGGSCATAGTCTCTCACTGCTTGGTTCGCGCGCTCGGGTGCCAGAGTGCACAGT 344
QY 180 GCCAACCCAGTCTGTGTGCAACCCGACCTGCTTCCCACTTCTCGTGGAGCGCCGAG 239
Db 345 GCCAACCCAGTCTGTGTGCAACCCGACCTGCTTCCCACTTCTCGTGGAGCGCCGAG 404
QY 240 GATGTGTCATGTGTCAGAAACAGCCAGTGTGTGTGCAAGCGCGTGCCTCCGCCACG 299
Db 405 GACGTGTACATTTGTCAGAAACAGCCAGTGTGTGTGCAAGCGTGTGCCTCCGCCACG 464
QY 300 CAGATCTTCTCAAGTCAACGGGGAGTGGTGCGCCAGTGGACCAAGTATCAGCGC 359
Db 465 CAGATCTTCTCAAGTCAACGGGGAGTGGGTTCGCGAGTGCATCAGTCAATTAAGCGC 524
QY 360 AGCACAGCGGAGCAGTGTGTGAGCCGACCATGGAGTCCGCATTAAATGTCTCAGGCG 419
Db 525 AGCACTGACGCGCAGCGGATTTGCCAACCATGAGGTCCGGATCAAGTATCAAGGCGAG 584
QY 420 CAGTCCAGAAAGTGTTCGGCTCGAGGAATPACTGGTCCAGTGTGGTCCATGGAGTCC 479
Db 585 CAGTCCAGAAAGTGTTCGGCTCGAGGAGTACTGGTCCAGTGTGTGGCATGGAGTCC 644
QY 480 TCGGGCCACCCAGAGTCAAGAGCTACATCCGATAGCCAGTTCGCGAAGAACTTC 539
Db 645 TCAGGAACCCACAAAGCCAGAGGCTTACATCCGGATTGCTATTTCGCGAAGAACTTT 704
QY 540 GAGCAGGAGCGCTGGCCAGAGAGTGTCCCTGGAGCAGGCGCATCGTGTGCCCTCGCGT 599
Db 705 GAGCAGGAGCGCTGGCCAGAGAGTGTCACTGGAGCAAGGCATTGTGTACCTTGTCCG 764
QY 600 CCACCGAGGCGCATCCTTCCAGCCGAGTGGAGTGGCTTCGGAACAGGAGCTGGTGGAC 659
Db 765 CCCCAGGAAGGAATCCCCCAGCTGAGTGGAGTGGCTTCGGAATAGGACCTCGTGGAC 824
QY 660 CCGTCCCTGAGCCCAATGTATACATCAGCGGAGCAGACCTGTGGTGGCAGAGGCC 719
Db 825 CCCTCCCTCGACCCCAATGTATACATCAGCGGAGCAGACCTGTGGTGGCAGAGGCC 884

QY 720 CGCTTGTGTGACACGGCCAACTACACTGCTGTGCGCAAGAAACATCGTGACCTCGCGC 779
Db 885 CGCTTGTGTGACACCTCCAACTACACTGCTGTGCGCAAGAAACATCGTGACCTCGCGC 944
QY 780 AGCGCTTGTGTGCTGT 839
Db 945 AGCGCTTGTGTGCGCGCGTCA----- 963
QY 840 TCCGTCTGTGAGCGCCAGCTGTGGCGCGCTGTGCGAGAAACGGAGCGGAGTGCACCAAC 899
Db 964 ----- 963
QY 900 CCGGCGCTCTTCAACGGGGGCGCTTTCTGTGAGGGGAGAGATGTCCATGACCGACCGTC 959
Db 964 ----- 963
QY 960 TCCTCTCTGT 1019
Db 964 -----TTGTTTATGT 1013
QY 1020 GGGCTGTGTGACACCACTGTGGCGGAGCGCTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 1079
Db 1014 GGGCTGTGTGACACCACTGTGGCGGAGCGGGAGTGTGTGTGTGTGTGTGTGTGTGTGTGT 1073
QY 1080 GGGGAGGAGTGCCAGGGCACTGACCTGTGACACCCGCAACTGTGTGTGTGTGTGTGTGT 1139
Db 1074 GGTGAGGAGTGTGGGGGT 1133
QY 1140 CACAGTGTCTTGTGGCGCTGAGGAGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 1199
Db 1134 CACAGTGTCTTGTGGCGCGGAGAGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 1193
QY 1200 TGCCTGT 1259
Db 1194 TGCCTGT 1253
QY 1260 GACTCAGATGT 1319
Db 1254 GACTCAGAGT 1313
QY 1320 CCCAGCAAGCAGACACACCCCATCTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 1376
Db 1314 CCCAGCAAGCAGACACATCCCATCTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 1373
QY 1377 ACCACCTACAGGGCAGTCTGT 1436
Db 1374 ACCACCTACAGGGCAGCTGT 1433
QY 1437 ACCAATGGGCACTGT 1496
Db 1434 TCTAATGGTCACTGT 1493
QY 1497 CCCACCTCTGAGGCGGAGGAGT 1556
Db 1494 CCCACCTCTGAGGCTGAGGAGT 1553
QY 1557 CTGCCCCGAGGCAACAGCAATGACCTATGGGACCTTCACTTCTGTGTGTGTGTGTGTGTGT 1616
Db 1554 CTGCCCCGCGGTACAGCAACATGGCCCTATGGGACCTTCACTTCTGTGTGTGTGTGTGTGT 1613
QY 1617 ATGATCCCTAATCAGATGTACGCTCTCTATCCCGCCAGATGCATATCCCGAGGGAAG 1676
Db 1614 ATGATCCCTAATCAGGAAATCAGCTCTCTATACCGCCGAGGCAATCCCGAGGGAAG 1673
QY 1677 ATCTATGAGATCTACCTCAGCTGCACAGCGGAGACGTGTGTGTGTGTGTGTGTGTGTGTGT 1736
Db 1674 ATCTAAGATCTACCTCAGCTGCACAGCGGAGACGTGTGTGTGTGTGTGTGTGTGTGTGT 1733
QY 1737 TGTGAGACCTGT 1796
Db 1734 TGTGAGACCTGT 1793

1797 CCAGTATCTCTGGCTATGGAACAATGTGGGAGAGCCAGCCCTGACAGCTGAGCTGGCG 1856
1794 CCAGTCTATCTCTGGCTATGGAACAATGTGGGAGAGCCAGCTGCGAGCTGGCGCTGGCG 1853
1857 CTCAAAAGAGAGTCTGTGGAGGAGCTGGGAGCAGAGTGTCTGCACCTGGCGGAGGAG 1916
1854 CTCAAAAGAGAGTCTGTGGAGGAGCTGGG---AGGACGTGTCTGCACCTGGTGGAGGAG 1910
1917 GCGCCCTCCACCTCTACTACTGCGAGCTGGAGGCGAGTGTCTGCTCTTCAACCGAG 1976
1911 TCGCCCTCTCATCTCTACTACTGCGAGCTGGAGGCGGCGCTGTATGTCTTTACCGAG 1970
1977 CAGTGGGCGCTTTGGCTGTGGAGAGGCGCTCAGGTGGCTGCGCCGAGGCGCTC 2036
1971 CAGTGGGCGCTTTGGCTGTGGAGAGGCGCTCAGGTGGCTGCGCCGAGGCGCTC 2030
2037 AAGTGTCTCTCTGTGGCGCGCTGCTGCACTCTCCCTCGAGTCAAAATCCGGGTCTAC 2096
2031 AGGTCTCTCTGTGGCGCGCTGCTGCACTCTCCCTCGAGTCAAAATCCGGGTCTAC 2090
2097 TGCCTGCATGACACCCAGATGCACTCAAGAGGTGGTGCAGCTGGAGAGCAGCTGGG 2156
2091 TGCCTGCATGACACCCAGATGCACTCAAGAGGTGGTGCAGCTGGAGAGCAGCTGGG 2150
2157 GGACAGCTGATCCAGAGCAGCGGTCTGCACTTCAAGGACAGTTACACACCTGGCG 2216
2151 GGACAGCTGATCCAGAGCAGCGGTCTGCACTTCAAGGACAGTTACACACCTGGCG 2210
2217 CTATCCATCCAGATGTCGCCAGCTCTCTGTGGAGAGTAAAGTCTCTGTGAGTACCAG 2276
2211 CTGTCCATCCAGATGTCGCCAGCTCTCTGTGGAGAGTAAAGTCTCTGTGAGTACCAG 2270
2277 GAGATCCCTTTTATCACATCTGGAATGCGACGCGGTACTGTGCACTGCACCTTACC 2336
2271 GAGATCCCTTTTATCACATCTGGAATGCGACGCGGTACTGTGCACTGCACCTTACC 2330
2337 CTGGAGCGTGTAGCCCGAGCTAGTGTGCTGCAAGCTGCAAGCTGGGTGGAGGTTG 2396
2331 CTGGAGCGTGTAGCCCGAGCTAGTGTGCTGCAAGCTGCAAGCTGGGTGGAGGTTG 2390
2397 GAGGCGGAGCGGAGAGCTTACAGATCAACTTCAACATCACCAGGACACAAAGTTGCT 2456
2391 GAGGAGATGGACAGAGCTTACAGATCAACTTCAACATCACCAGGACACAAAGTTGCT 2450
2457 GAGTCTGCTGCTGAGAGTGAAGCGGGGTCCAGCCCTGTGTGGGCCCGAGTCTTC 2516
2451 GAAATGCTGCTGAGAGTGAAGCGGGGTCCAGCCCTGTGTGGGCCCGAGTCTTC 2510
2517 AAGATCCCTTCTCTATTGGCAGAGATAATTTCCAGCTGGACCCAGCCCTGTAGCGG 2576
2511 AAGATCCCTTCTCTATTGGCAGAGATAATTTCCAGCTGGACCCAGCCCTGTAGCGG 2570
2577 GGTGCGGAGTGGCGGACTCTGGCCGAGAACTCCAGCTGGACAGCCATCTCAGCTTCTTT 2636
2571 GGTGCGGAGTGGCGGACTCTGGCCGAGAACTCCAGCTGGACAGCCATCTCAGCTTCTTT 2630
2637 GGTGCGGAGTGGCGGACTCTGGCCGAGAACTCCAGCTGGACAGCCATCTCAGCTTCTTT 2696
2631 GGTGCGGAGTGGCGGACTCTGGCCGAGAACTCCAGCTGGACAGCCATCTCAGCTTCTTT 2690
2697 AACGGCAACTCTGAGCAGCTGGTGGAGCAGTGGTGGTGGGCGGAGCAGCTGGC 2756
2691 AACGGCAACTCTGAGCAGCTGGTGGAGCAGTGGTGGTGGGCGGAGCAGCTGGC 2750
2757 CTCTTCAAGTGTGGAGGCTGAGTGTGAGGCGGCGAGCCCGAGCA--CTTACACTCT 2814
2751 CTCTTCAAGTGTGGAGGCTGAGTGTGAGGCGGCGAGCCCGAGCA--CTTACACTCT 2810
2815 CACAGCTTTGGACCCACCAAGAGCAGGAGAGCCGAGGCGCTTCCCGACACC 2874
2811 CACAGCTTTGGACCCACCAAGAGCAGGAGAGCCGAGGCGCTTCCCGACACC 2870
2875 GGGGAGA 2881

Db 2871 CCGGGGA 2877
RESULT 10
BC009333
LOCUS
DEFINITION
ACCESSION
VERSION
KEYWORDS
SOURCE
ORGANISM
REFERENCE
AUTHORS
TITLE
JOURNAL
REFERENCE
AUTHORS
TITLE
JOURNAL
REMARK
COMMENT
FEATURES
source

BC009333 2688 bp mRNA linear PRI 19-DEC-2003
Homo sapiens cDNA clone IMAGE:4126760, partial cds.
BC009333
BC009333.2 GI:40226527
Homo sapiens (human)

Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
1 (bases 1 to 2688)
Strausberg, R.L., Feingold, E.A., Grouse, L.H., Derge, J.G.,
Altschul, R.D., Collins, F.S., Wagner, L., Shenmen, C.M., Schuler, G.D.,
Hopkins, R.F., Jordan, H., Moore, T., Max, S.I., Wang, J., Hsieh, P.,
Diatchenko, L., Marusina, K., Farmer, A.A., Rubin, G.M., Hong, L.,
Stapleton, M., Soares, M.B., Bonaldo, M.P., Casavant, T.L.,
Scheetz, T.E., Brownstein, M.J., Usdin, T.B., Lohman, P.J.,
Abramson, R.D., Mullikin, S.J., Bosak, S.A., McEwan, P.J.,
McKernan, K.J., Malek, J.A., Gunaratne, P.H., Richards, S.,
Worley, K.C., Hale, S., Garcia, A.M., Gay, L.J., Hulyk, S.W.,
Villalón, D.K., Muzny, D.M., Sodergren, E.J., Lu, X., Gibbs, R.A.,
Fahey, J., Helton, E., Kettner, M., Madan, A., Rodriguez, S.,
Sanchez, A., Whitting, M., Madan, A., Young, A.C., Shevchenko, Y.,
Bouffard, G.G., Blakesley, R.W., Touchman, J.W., Green, E.D.,
Dickson, M.C., Rodriguez, A.C., Grimwood, J., Schmutz, J., Myers, R.M.,
Butterfield, Y.S., Krzyzanski, M.I., Skalska, U., Smal, M.A.,
Schnerch, A., Schein, J.E., Jones, S.J. and Marra, M.A.
Generation and initial analysis of more than 15,000 full-length
human and mouse cDNA sequences
Proc. Natl. Acad. Sci. U.S.A. 99 (26), 16899-16903 (2002)
2 (bases 1 to 2688)
Strausberg, R.
Direct Submission
Submitted (12-JUN-2001) National Institutes of Health, Mammalian
Gene Collection (MGC), Cancer Genomics Office, National Cancer
Institute, 31 Center Drive, Room 11A03, Bethesda, MD 20892-2590,
USA

NIH-MGC Project URL: <http://mgc.nci.nih.gov>
On Dec 19, 2003 this sequence version replaced gi:14424611.
Contact: MGC help desk
Email: cgabbs@mail.nih.gov
Tissue procurement: ATCC
cDNA Library Preparation: Rubin Laboratory
cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
DNA Sequencing by: National Institutes of Health Intramural
Sequencing Center (NISC),
Gaithersburg, Maryland;
Web site: <http://www.nisc.nih.gov/>
Contact: nisc.mgc@nih.gov
Akhter, N., Ayala, K., Beckstrom-Sternberg, S.M., Benjamin, B.,
Blakesley, R.W., Bouffard, G.G., Breen, K., Brinkley, C., Brooks, S.,
Dieckman, N.D., Granite, S., Guan, X., Gupta, J., Haghighi, P.,
Hansen, N., Ho, S.-L., Karlins, E., Kwong, P., Lari, P., Legaspi, R.,
Maduro, Q.L., Masfello, C., Maskeri, B., Mastrian, S.D., McCloskey, J.C.,
McDowell, J., Pearson, R., Stantripop, S., Thomas, P.J., Touchman, J.W.,
Tsougeon, C., Vogt, J.L., Walker, M.A., Wetherby, K.D., Wiggins, L.,
Young, A., Zhang, L.-H. and Green, E.D.

Clone distribution: MGC clone distribution information can be found
through the I.M.A.G.E. Consortium/LLNL at: <http://image.llnl.gov>
Series: IRAL Plate: 26 Row: g Column: 22.

Location/Qualifiers
1. 2688
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"

1757	CATCGTTAGCTGTGGACCCCTCGCGTCTGCTCACCCGGCCAGTCACTCTGGCTATGGA	1816
601	CATCGTTAGCTGTGGACCCCTCGCGTCTGCTCACCCGGCCAGTCACTCTGGCTATGGA	660
1817	CCACTGTGGGGAGCCCAAGCCCTGACAGCTGGAGCTCGGCTCAAAAGCAGTCTGGCGGA	1876
661	CCACTGTGGGGAGCCCAAGCCCTGACAGCTGGAGCTCGGCTCAAAAGCAGTCTGGCGGA	720
1877	GGGCAAGCTGGGAGCAGGATGTGTGCACCTGGGCGAGGAGCGCCCTCCCACTCTACTA	1936
721	GGGCAAGCTGGGAGCAGGATGTGTGCACCTGGGCGAGGAGCGCCCTCCCACTCTACTA	777
1937	CTGCGAGCTGGAGGCGAGTGCCTGCTACGTCTTCAACGAGCAGCTGGGCGCTTTGCCCT	1996
778	CTGCGAGCTGGAGGCGAGTGCCTGCTACGTCTTCAACGAGCAGCTGGGCGCTTTGCCCT	837
1997	GGTGGGAGAGGCGCCTCAGCGTGGCTGCGGCCAAGCGCTCAAGCTGCTCTGTTTGGGCG	2056
838	GGTGGGAGAGGCGCCTCAGCGTGGCTGCGGCCAAGCGCTCAAGCTGCTCTGTTTGGGCG	897
2057	GGTGGGCTGCACCTCCCTCGAGTACCAACTCCGGGTCTACTGTGCTCATGACACCCACGA	2116
898	GGTGGGCTGCACCTCCCTCGAGTACCAACTCCGGGTCTACTGTGCTCATGACACCCACGA	957
2117	TGCACCTCAAGGAGGTGGTGCAGCTGGAGAGCAGCTGGGGGGHACAGTGCATCAGAGCC	2176
958	TGCACCTCAAGGAGGTGGTGCAGCTGGAGAGCAGCTGGGGGGHACAGTGCATCAGAGCC	1017
2177	ACGGGTCTTGCACTTCAAGGACAGTTTACCACACCTCGCCCTATCCATCCAGATGTGCC	2236
1018	ACGGGTCTTGCACTTCAAGGACAGTTTACCACACCTCGCCCTATCCATCCAGATGTGCC	1077
2237	CAGCTCCCTGTGGAAGAGTAAGCTCCTGTGAGCTACCAAGAGATGCCCTTTTATCACAT	2296
1078	CAGCTCCCTGTGGAAGAGTAAGCTCCTGTGAGCTACCAAGAGATGCCCTTTTATCACAT	1137
2297	CTGGAATGGCACGACGCGGTACTTGCACCTGGCACTTCAACCTGGAGCGTGTGAGAGG	2356
1138	CTGGAATGGCACGACGCGGTACTTGCACCTGGCACTTCAACCTGGAGCGTGTGAGAGG	1197
2357	CACAGTGCACCTGGCTGCAAGCTGTGGGTGGGAGGTGGAGGCGAGCGGACAGCTT	2416
1198	CACAGTGCACCTGGCTGCAAGCTGTGGGTGGGAGGTGGAGGCGAGCGGACAGCTT	1257
2417	CAGCATCAACTTCAACATCACAAGGACACAAGGTTTGTGAGCTGCTGGCTCTGGAGAG	2476
1258	CAGCATCAACTTCAACATCACAAGGACACAAGGTTTGTGAGCTGCTGGCTCTGGAGAG	1317
2477	TGAAGCGGGGTCCAGCCCTGGTGGGCGCCAGTGCCTTCAAGATCCCTTCCCTCATTCG	2536
1318	TGAAGCGGGGTCCAGCCCTGGTGGGCGCCAGTGCCTTCAAGATCCCTTCCCTCATTCG	1377
2537	GCAGAAGATATTTCCAGCTGGACCCACCTGTAGGGGGGTGCCGACTGGCGACTCT	2596
1378	GCAGAAGATATTTCCAGCTGGACCCACCTGTAGGGGGGTGCCGACTGGCGACTCT	1437
2597	GGCCGAGAACTCCACCTGGACAGCCATCTCAGCTTCTTTTGCTCCAAAGCCCGAGCCAC	2656
1438	GGCCGAGAACTCCACCTGGACAGCCATCTCAGCTTCTTTTGCTCCAAAGCCCGAGCCAC	1497
2657	AGCCATGATCTCAACCTGTGGGAGCGCGGCACTTCCCAACAGGCAACCTCAGCCAGCT	2716
1498	AGCCATGATCTCAACCTGTGGGAGCGCGGCACTTCCCAACAGGCAACCTCAGCCAGCT	1557
2717	GGCTCAGCAGTGGCTGGACTGGGCGACGAGAGCTGTGAGTGTGGAGGC	2776
1558	GGCTCAGCAGTGGCTGGACTGGGCGACGAGAGCTGTGAGTGTGGAGGC	1617
2777	TGAGTGTGAGGCGCGGCGAGCCGACCACTTCACTGTGAGTGTGGAGGC	2836
1618	TGAGTGTGAGGCGCGGCGAGCCGACCACTTCACTGTGAGTGTGGAGGC	1677


```
QY 2670 AACCTGTGGAGGCGCGGCACTTCCCAACGGCAACCTCAGCCAGCTGGCTGCAGCAGTG 2729
Db 1552 AACCTGTGGAGGCGCGGCACTTCCCAACGGCAACCTCAGCCAGCTGGCTGCAGCAGTG 1611
QY 2730 GCTGGACTGGGCGCAGCAGAGCTGGCTC-TTCACAGTG-TGAGAGGCTGAGTGTGAG 2787
Db 1612 GCTGGAGACTGGCAGCAGAGCAGTGGCTTCTTTACAGTGTTCGGAGGCTGAGTGTGAG 1671
QY 2788 GCCGGCAGGCGCGCACACCTACACTCTCACAGCTTTGGCA--CCCAACCAAGCAGAGCA 2845
Db 1672 GCCGGCCAGG--CGAACACTAGCAATTTACAGTTTTGGGAACCCACCAGAGGACAGGCA 1729
QY 2846 GAAGCCGGACAGGGGCGCTTCCCAACACCGGGGAGA 2881
Db 1730 GAAGCCGGACAGGGGCTTTTCCCAAAACCGGGGAGA 1765

RESULT 12
AX054976
LOCUS
DEFINITION Sequence 91 from Patent WO0073328.
ACCESSION AX054976
VERSION AX054976.1 GI:12228344
KEYWORDS synthetic construct
SOURCE synthetic construct
ORGANISM artificial sequences.
REFERENCE
1 van Crielinge,W., Roelens,I., Bogaert,T. and Verwaerde,P.
Unc-5 constructs and screening methods
Patent: WO 0073328-A 91 07-DEC-2000;
Devgen NV (BE)
FEATURES
Location/Qualifiers
1..9700
/morganism="synthetic construct"
/mol_type="unassigned DNA"
/db_xref="taxon:32630"
/note="plasmid pGC1037"
ORIGIN
Query Match 44.8%; Score 1291.2; DB 6; Length 9700;
Best Local Similarity 99.5%; Pred. No. 9e-204;
Matches 1306; Conservative 0; Mismatches 3; Indels 3; Gaps 1;
QY 1478 CACACTGCACACAGCTCTCCACCTCTGAGCCCGAGGAGTTGCTCCCGCTCTCCAC 1537
Db 7 CACACTGCACACAGCTCTCCACCTCTGAGCCCGAGGAGTTGCTCCCGCTCTCCAC 66
QY 1538 CCAGAACTACTTCCGCTCCCTGCTCCCGGAGGACCAACATGACCTATGGACCTTCAA 1597
Db 67 CCAGAACTACTTCCGCTCCCTGCTCCCGGAGGACCAACATGACCTATGGACCTTCAA 126
QY 1598 CTTCTCGGGGCGCGGTGATGATCCCTAATACAGGTATCAGCTTCCCTATCCCGCCAGA 1657
Db 127 CTTCTCGGGGCGCGGTGATGATCCCTAATACAGGTATCAGCTTCCCTATCCCGCCAGA 186
QY 1658 TGCCATACCCCGAGGAAAGATCTATGAGATCTACCTCAGCTGCACAAAGCCGGAAGAGT 1717
Db 187 TGCCATACCCCGAGGAAAGATCTATGAGATCTACCTCAGCTGCACAAAGCCGGAAGAGT 246
QY 1718 GAGTTGCCCTAGTGGCTGTAGACCTGTGAGTCCCATCGTTAGCTGTGGACCCCC 1777
Db 247 GAGTTGCCCTAGTGGCTGTAGACCTGTGAGTCCCATCGTTAGCTGTGGACCCCC 306
QY 1778 TGGCGTCTGTCTACCCGGCCAGTATCTCTGGCTATGACCACTGTGGGAGCCAGCCC 1837
Db 307 TGGCGTCTGTCTACCCGGCCAGTATCTCTGGCTATGACCACTGTGGGAGCCAGCCC 366
QY 1838 TGACAGCTGGAGCTGGCGCTCAAAAGACAGTCTGTGGAGGGCAGCTGGGAGAGGAGTGT 1897
Db 367 TGACAGCTGGAGCTGGCGCTCAAAAGACAGTCTGTGGAGGGCAGCTGGG---AGGATGT 423
QY 1898 GCTGCACCTGGGCGAGGAGGCGCCCTCCACCTCTACTTACTGTCCAGCTGGAGGCCAGTGC 1957
```

```
Db 424 GCTGCACCTGGGCGAGGAGCGCCCTCCACCTCTACTGTCAGCTGGAGGCCAGTGC 483
QY 1958 CTGCTACGTTCTTACCGAGCAGCTGGGCGCGCTTGGCCCTGGTGGGAGAGCCCTCAGCGT 2017
Db 484 CTGCTACGTTCTTACCGAGCAGCTGGGCGCGCTTGGCCCTGGTGGGAGAGCCCTCAGCGT 543
QY 2018 GGTGCGCCCAAGCGCTCAAGCTGCTTCTTGTGGCGCGGTGGCTGCACTCCCTCGA 2077
Db 544 GGTGCGCCCAAGCGCTCAAGCTGCTTCTTGTGGCGCGGTGGCTGCACTCCCTCGA 603
QY 2078 GTACAACTCCCGGCTTACTGCTGCTGATGACACCCAGATGCACTCAAGAGGTGGTGA 2137
Db 604 GTACAACTCCCGGCTTACTGCTGCTGATGACACCCAGATGCACTCAAGAGGTGGTGA 663
QY 2138 GCTGGAGAAAGCAGCTGGGGGGACAGCTGATCCAGGAGCCACGGTCTCTGCACTTCAAGGA 2197
Db 664 GCTGGAGAAAGCAGCTGGGGGGACAGCTGATCCAGGAGCCACGGTCTCTGCACTTCAAGGA 723
QY 2198 CAGTTACCAACAGCTGGGCTATCCATCCAGATGTGCCAGCTCCCTGTGGAAGATGA 2257
Db 724 CAGTTACCAACAGCTGGGCTATCCATCCAGATGTGCCAGCTCCCTGTGGAAGATGA 763
QY 2258 GCTCCTTGTGACCTACCGAGATCCCTTTTATCACATCTGGAATGGACGAGCGTA 2317
Db 784 GCTCCTTGTGACCTACCGAGATCCCTTTTATCACATCTGGAATGGACGAGCGTA 843
QY 2318 CTTGCACTGCACCTTCACTGAGGCTGTGAGCCCGCAGCACTAGTACCTGGCTGCA 2377
Db 844 CTTGCACTGCACCTTCACTGAGGCTGTGAGCCCGCAGCACTAGTACCTGGCTGCA 903
QY 2378 GCTGTGGTGTGGAGGTGGAGGGGACGGGCGAGAGCTTTCAGCATCAACTTCAACATCAC 2437
Db 904 GCTGTGGTGTGGAGGTGGAGGGGACGGGCGAGAGCTTTCAGCATCAACTTCAACATCAC 963
QY 2438 CAAGGACACAGGTTTGTGCTGAGCTGCTGCTGCTGAGAGTGAAGGGGGTCCGAGCCT 2497
Db 964 CAAGGACACAGGTTTGTGCTGAGCTGCTGCTGCTGAGAGTGAAGGGGGTCCGAGCCT 1023
QY 2498 GGTGGGCCCCAGTGCCTTCAAGATCCCTTCTCATTCGGCAGAGAGATAATTTCCAGCCT 2557
Db 1024 GGTGGGCCCCAGTGCCTTCAAGATCCCTTCTCATTCGGCAGAGAGATAATTTCCAGCCT 1083
QY 2558 GGACCCACCTGTAGGGGGGTGCGACTGCGGGACTCTGGCCCAAGAACTCCACTGGA 2617
Db 1084 GGACCCACCTGTAGGGGGGTGCGACTGCGGGACTCTGGCCCAAGAACTCCACTGGA 1143
QY 2618 CAGCCATCTCAGCTTCTTTGCTTCCAGCCCGAGCCGACAGCCATGATCTCAACTGTG 2677
Db 1144 CAGCCATCTCAGCTTCTTTGCTTCCAGCCCGAGCCGACAGCCATGATCTCAACTGTG 1203
QY 2678 GGAGCGCGCACTTCCCAACGGCACTCAGCAGCTGGCTGCAGCAGTGGCTGGACT 2737
Db 1204 GGAGCGCGCACTTCCCAACGGCACTCAGCAGCTGGCTGCAGCAGTGGCTGGACT 1263
QY 2738 GGGCCAGCAGACCTGCGCTCTTTCAGTGTGCGAGGCTGAGTGTGAGGC 2789
Db 1264 GGGCCAGCAGACCTGCGCTCTTTCAGTGTGCGAGGCTGAGTGTGAGTC 1315
```

```
RESULT 13
MMU72634
LOCUS
DEFINITION Mus musculus rostral cerebellar malformation protein (rcm) mRNA, complete cds.
ACCESSION U72634
VERSION U72634.1 GI:2088526
KEYWORDS Mus musculus (house mouse)
SOURCE Mus musculus
ORGANISM Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
REFERENCE 1 (bases 1 to 9299)
```

AUTHORS Ackerman,S.L., Kozak,L.P., Przyborski,S.A., Rund,L.A., Boyer,B.B. and Knowles,B.B.
 TITLE The mouse rostral cerebellar malformation gene encodes an UNC-5-like protein
 JOURNAL Nature 386 (6627), 838-842 (1997)
 MEDLINE 97271898
 PUBMED 9126743
 REFERENCE 2 (bases 1 to 9299)
 AUTHORS Ackerman,S.L., Kozak,L.P., Rund,L.A. and Knowles,B.B.
 TITLE Direct Submission
 JOURNAL Submitted (25-SEP-1996) The Jackson Laboratory, 600 Main Street, Bar Harbor, ME 04609, USA
 FEATURES
 Location/Qualifiers
 1..9299
 /organism="Mus musculus"
 /mol_type="mRNA"
 /strain="C57B6/SJL"
 /db_xref="taxon:10090"
 /chromosome="3"
 /map="between Amy1 and Mpmv9"
 1..9299
 /gene="rcm"
 151..2946
 /gene="rcm"

gene

CDS

/function="transmembrane receptor"
 /note="similar to C. elegans UNC-5"
 /codon_start=1
 /product="rostral cerebellar malformation protein"
 /protein_id="AAB54103.1"
 /db_xref="GI:2088527"
 /translation="MRKGLRATARRCGGLGLVLLQMLVLPALALLSASGTGSAADDE
 FPHLEPFPDPPELPHLEPEAYIVKKNPNVLYCKAPATQIYFKNSRWHQ
 KDHVDERVDSGLIVREVSEIIRQVEELFGPEDEYWCQVAVSSAGTTKSRKAYV
 RIAYLRKPEQPLEQLEVLQCRPEGIPVAEVLKNEIDIDPAEDRFYI
 TIDNLIQKARLSDTANTYCAKNIIVAKRKSTATVIVVANGWSTWTSVCSNRC
 GRGYKRYFTCNPLNGAFEGOSVOKIATLCPVDGRWTSWKSCTGCTECLA
 WRRRECTAPKNGKDCGLVLQSKNCTDGLCMQAPDSDVALVGVIVATVCLIA
 ITVVVALFVKNHRDFESDIIDSSALNGDFQVNIKAARQDLAVPPDLTAAAMRY
 GPVALHDVSDKIPTNLSPIPLPNLKIKVNSGAVTPQDDLAFFSKLSPQMTQS
 LLEENALINKLSLARQDPSCTAGTNSLGHLLIPNSGVSLIPAGAIQGRVYE
 MVTYVHRKNRPMEDSQTLLTPVSCGPPGALLTRPVILTHHCADSPEDTKIQL
 KNOAVOGEDVVVVEENFTPCYIQLDAEACHILTENLSVALVQSTTKAAAKRL
 KLAIFGLCSLSIYSIRVYCLDDTDQALKEVLOROMGOLLEPKALHFKGLTHN
 LRLSHDIAHLSWKLAKYQIPIFYHWSGQRLNHCFTLERLSINTVELVKLC
 VQVEGEGQIFQNLTVSEPTGIDILDPASTITVTGSPAFSIPLEPIRKLCSLL
 DAPQTRGHDWRMLAHKLNLDRLNYEATKSSPTGVILDLWEAQNFDPGNLSMLAAVLE
 EMGRHETVTVSLAEGGY"

ORIGIN

Query Match 32.6%; Score 940.2; DB 10; Length 9299;
 Best Local Similarity 61.9%; Pred. No. 9,9e-146;
 Matches 1642; Conservative 0; Mismatches 943; Indels 66; Gaps 7;

QY 190 TGCCTGGTCCAAACCGGACCTGCTCCCACTTCCTGGTGGAGCCCGGAGATGTGTACA 249
 Db 308 TTCCATCTGACCACTGAGCATTGCCACACTTCCTCATTTAGCCCGAGAGCTTACA 367
 QY 250 TCCTCAAGAACAAAGCCAGTCTGCTGTGTGTGCAAGGCCGCTGCCCGCCACAGACTTCT 309
 Db 368 TTGTGAAGAACAAAGCCAGTCTGAACTGTATTTGTAAGCCAGCCCTGCCACCCAGACTTACT 427
 QY 310 TCAAGTGCACGGGAGTGGTGGCGCAGGTGGACACAGCTGATCGAGCGGACGACACAGC 369
 Db 428 TCAAGTGCACGGCAGTGGGTTTCATCAGAAGGACCCAGTAGTAGACGAGAGTAGATG 487
 QY 370 GGAGCAGTGTGAGCCGACCATGAGGTCCGCAATTAATGTCTCAAGGCAGCAGGTTCGAGA 429
 Db 488 AATCCTCTGGTCTAATTTGTAGAGAAAGTGGACATTGAGATTCACGCCAGCAGGTGGAGG 547
 QY 430 AGGTGTTTGGGCTGGAGAAATATCTGGTCCAGTCCGCTGGCATGGAGTCTCTCGGGACCA 489
 Db 548 AACTGTTTGGGCTGAAGATTAATCTGGTGCCAGTGTGTGGCTGGAGCTCAGCAGGACCTA 607

490 CCAAGAGTCAGAGCGCTTACATCCGATAGCCAGATTGCGCAAGAACTTCAGAGCAGGAGC 549
 Db 608 CGAAGATCGAAGGACATACGTGCGCATTTGGTATCTGCGAAGACATTCGAGCAGGAAC 667
 QY 550 CGCTGGCCAAAGGAGTGTCTCTGAGCAGAGGCAATGCTGCCCTGCCCTCCACCGAGG 609
 Db 668 CTTGGGAAAGGAAGTGTCTTTGGAGCAGGAAGTCTTACTCCAGTGTCTGGCCACCTGAAG 727
 QY 610 GCATCCCTCCAGCCGAGGTGGAGTGGCTCCGGAACGAGAGACCTGGTGACCCGCTCCCTGG 669
 Db 728 GGATCCCAAGTGGCTGAGGTGAATGGCTAAAGAAATGAAGACATAATTTGATCTCTGTGAAG 787
 QY 670 ACCCCAAATGTATACATCACGGGGAGCAGACGCTGGTGGCGACAGGCGCCGCTGTGCTG 729
 Db 788 ATCGGAACTTTTATATTACTATCGATCACACCTGATCATCAAGCAAGCCGCACTCTCAG 847
 QY 730 ACACGGCCAACTACACCTGCTGGCGCAAGAACATCGTGGCAGCTGCGCCGACGGCTCCG 789
 Db 848 ATACAGCAAAATATACCTGTGTGGCAAAAATATTGTGCCAAGAGAAAAGCACACAG 907
 QY 790 CTGCTGTCTCATCTCTACGTGAACGGTGGTGTGCGACGTGGACCGAGTGGTCCGCTGCA 849
 Db 908 CCACCTGTCTATCTGTATTTAATGTTGCTGTGCTCCACCTGGACAGAGTGTCTGTGTA 967
 QY 850 GCGCCAGCTGTGGCGCGCTCGCAGAAAACGAGCGGAGCTGCACCAACCCGCGGCGCTC 909
 Db 968 ACAGCGCGTGTGGCGGAGGATATCAGAAACGCAAGAACCTGCACCAACCCGAGCCAC 1027
 QY 910 TCAACGGGGCGCTTTCTGTGAGGGGAGAAATGTCCATGACCGCACCGCTCTCTCTCTGC 969
 Db 1028 TCAATGTTGGGCGCTTCTGTGAGGGGAGAGTGTGCAGAA---AATAGCATGCATACCT 1084
 QY 970 TTGCTCTGTGACGCGACCTCGAGCCCTGGAGCAAGTGGTGGCGCTCTGGGCTCGACT 1029
 Db 1085 TATGTCAGTGTATGTAGTGAGCTTCAITGAGCAAAATGGTCAACTGTGGAGTGAAT 1144
 QY 1030 GCACCCACTGTGGCGAGCCCTGTGCTGTGACCCAGCAGCCCGCAACCGAGGGAGGAGT 1089
 Db 1145 GCACCCACTGTGGCGAGGGAGTGTACAGCAGCAGCCCGCAAGAACCGGGGTAAAGACT 1204
 QY 1090 GCAGGGCACTGACTGTGAGACACCGCAACTGTACAGTCACTCTGTGTACAGTGCCT 1149
 Db 1205 GTGATGGCTGTCTCTCAATCCAAAGACTGCACTGATGGGTGTGCACTGAGCGTGTCT 1264
 QY 1150 CTGGCCCTCAGAGCGTGGCCCTCTATGTGG---CCTCATCGCGCTGGCGCTGTGCTGG 1206
 Db 1265 CTGACTCAGATGATGTGGCTCTCTACGTGGGATTTGTATCGTGTAAACAGTCTGTCTGG 1324
 QY 1207 TCCTGTCTGTCTTGTCTCATCTCTTTATTGCGGGAAGAGGAGGGGCTGGACTCAG 1266
 Db 1325 CGATCACTGTGTGGTGGCCCTGTTTGTGTATCGGAAGAACCCCGTGACTTTGAGTCTG 1384
 QY 1267 ATGTGGCTGACTGTCTCATCTTCACTCAGGCTTCAGCCCGTCCAGCATCAAGCCGACGA 1326
 Db 1385 ACATCATTTGACTCTCTCAGCACTCAATGGGCGCTTTTCCAGCTGTGAACATCAAGCTGCCA 1444
 QY 1327 AAGCAGACAAACCCCATCTGCTCACCATCCAGCGGACCTTCAGCACCACCACTTACC 1386
 Db 1445 GACAAGATCTCTGGGTGTCTCCCTCTGACC-----TCACCTCAGTGCAGCATGTATCA 1498
 QY 1387 AGGCGAGTCTCTGTCCCGGAGGATG-----GGCCGAGCCCAAGTTCCAGTCCACCA 1441
 Db 1499 GGGGACCTGTCTATGCTCTCATGATGTCTCAGACAAAATCCCAATGACCAACTCTCAA 1558
 QY 1442 TGGGCACTGCTCAGCCCTCTGGTGGCGGCCCAACACTGACCACTGACAGTCTCTCCAC 1501
 Db 1559 TCTGGACCACTACCCAACTTGAATCAAAATGATCAAGCTCAGGTGTCTGCTCCTC 1618
 QY 1502 CTC-----TGAGCGCGAGGAGTGTGCTCTCCCGCTCTCTCAACC----- 1539
 Db 1619 CTAGGATGACCTTGGCGAGTCTCTCATCCAACTGTCAACCCAGATGATCAGTCTCTTGC 1678
 QY 1540 -----AGAACTACTTCCGCTCTCCCTCTCCCTGCGCCGAGGACCAAGCA 1575

Db 1679 TAGAGAAATGAGGCGCTTAACTGGAAGAACAGAGCCCTGCGAAGACAGACTGACCCATCCT 1738
 QY 1576 ACATGACCTATGGAGCCTTCAACTTCTCGGGGCGCGGTGATGATCCTTAATACAGGTA 1635
 Db 1739 GCACAGATTGTGTACCTTCAACTCTCTTGGGGGTACCTCATCATCTTCTTAATTCAGGAG 1798
 QY 1636 TCAGCCTCTCTATCCCCCAGATGCCATACCCCGAGGGAAGATCTATGAGATCTACCTCA 1695
 Db 1799 TAAGCTTGCTGATTCCCGCTGGGCGCAATCTCTCAGGGAGAGTCTATGAATGTATGTGA 1858
 QY 1696 CGCTGCACAGCCGGAAGAGCTGAGGTTGCCCTAGCTGCTGTGACAGCCCTGCTGAGTC 1755
 Db 1859 CTGTACACAGGAAGAAATATGAGGCCCGCCCATGGAAGACTCTCAGACCCCTACTTACC 1918
 QY 1756 CCATCGTTAGCTGTGACCCCTGGGCTGCTGCTCAACCCGCGGAGTCTGCTGGCTATGG 1815
 Db 1919 CTGTGTGAGCTGTGGGCTCTCTGGAGCTCTGTGACCCCGCTGTCTATCTCACTCTGC 1978
 QY 1816 ACCACTGTGGGAGCCAGCCCTGACAGCTGGAGCCTGCGCCTCAAAAAGCAGTCTGTGCG 1875
 Db 1979 ATCACTGTGCAGACCCAGCAGGAGCTGGAAGATCCAGCTCAAAAACAGGAGTGC 2038
 QY 1876 AGGAGCTGGGAGAGGATGTCTGCACCTGGCGAGGAGGCGCCCTCCACCTCTACT 1935
 Db 2039 AGGACAATGGGAGGATGTTGTGGTG--GTTGGGAGGAGAACTTCAACAACCCCTGTT 2095
 QY 1936 ACTGCCAGCTGGAGGCGAGTGCCTGCTAGCTCTTCCAGGAGCAGTGGGCGCTTTGCC 1995
 Db 2096 ACATTGAGCTGATGAGAGGCTTGCCATCTCTACAGAGAACCTCAGTACCTATGCGC 2155
 QY 1996 TGGTGGAGAGCCCTCAGCGTGGCTGCGCGCAAGCGCCTCAAGCTGCTTCTGTTGGCG 2055
 Db 2156 TGGTTGGGAGTCCACCAAGCAGCTGCCAAGGCTCTTAACCTGGCCACTTTGGCG 2215
 QY 2056 CGGTGCTGCACTCTCTCGAGTCAACATCCGGGTCTACTGCTGCAATGACACCCAG 2115
 Db 2216 CCCTCTGCTGCTCTCCCTGAGTACAGCATTAGCTCTACTGCTGATGACACAGG 2275
 QY 2116 ATGCACTAAGAGGTGTGAGCTGGAGGAGCAAGCAGCTGGGGGAGCAGTGTCCAGGAGC 2175
 Db 2276 ATGCTCTGAGGAAGTTCTACAACTGGAGGAGCAATGGGAGCAGCTCTTAGAAGAAC 2335
 QY 2176 CACGGTCTGCACTTCAAGGAGTACCAACACCTGCGCTATCCATCCACGATGTC 2235
 Db 2336 CCAAGCTCTTCAATTTAAGGAGCAGATCCACAACTGCGCTGTCTATTATGATGATGC 2395
 QY 2236 CCAGTCTCTGGAAGAGTAACTCTTGTGCTGAGCTACAGAGATCCCTTTTATCA 2295
 Db 2396 CCATTTCCCTCTGGAAGAGCAATTTGCTGGTAAAGTATCAGAAATTCATTTTACCACA 2455
 QY 2296 TCTGGAATGGACGAGCGGTACTTGTCACTGACCTTCACTGAGCGGTGTGACGCCCA 2355
 Db 2456 TCTGAGTGGCTCTCAAGAACCTTCACTGACCTTCACTGGAAGAGCTCAACCCCTAA 2515
 QY 2356 GCACTAGTGAOCTGCGCTGCAAGCTGTGGGTGTGGAGTGGAGGCGGAGGAGAGCT 2415
 Db 2516 ACACAGTGAACCTGTTTGAACCTCTGTGTGCGCAGGTTGAAGAGAGGAGGAGATCT 2575
 QY 2416 TCAGATCAACTTCAACATCACAAGAGCACAAGGTTTGTGAGTGTGCTGCTGTGAGA 2475
 Db 2576 TCCAGCTCAACTGTACTGTGTGAGAGGACCTTCTGCGATCGACTTCTCTCTGGACC 2635
 QY 2476 GTGAAGCGGGGTCCAGCCCTGTTGGGCGCCAGTGCCTTCAAGATCCCTCTCTCAATC 2535
 Db 2636 CTGCTAGTACCATCACTGTACCGGAGCAAGTCTTTCAGCAATCTCTCTCCCTATCC 2695
 QY 2536 GCGAAGAGATATTTCCAGCCTGGAGCCACCCCTGTAGCGGGGTCCGAGCTGGGAGTCT 2595
 Db 2696 GCGAAGCTATGACAGCCTGGATGCTCCCTCAACACAGAGGCCATGACTGTGGAGTGC 2755
 QY 2596 TGGCCCAAGAACTCACTGAGACGCCATCTAGCTTCTTGTGCTTCAAGCCGAGCCCA 2655

Db 2756 TGGCCCATAACTCAACCTCGACAGGTACTTGAATTACTTTGCCACCAATCGAGCCCA 2815
 QY 2656 CAGCCATGATCTCAACCTGTGGAGCGCGGACCTTCCCAACCGCAACCTCAGGCAGC 2715
 Db 2816 CTGGCGTAATCTCTGATCTTTTGGAGGACACAGAACTTCCAGATGGAACTCAGCATGC 2875
 QY 2716 TGGCTGCAGCAGTGGCTGAGCTGGCCAGCAGCAGCTGGCCCTTTCACAGTGTGCGAGG 2775
 Db 2876 TGGCAGCCGCTCTGGAAGAAATGGGAAGACATGAGACAGTGGTCTCTTGGCAGCAGA 2935
 QY 2776 CTGAGTGTGA 2786
 Db 2936 GACAGTATTGA 2946

RESULT 14
 AY187310
 LOCUS
 DEFINITION
 ACCESSION
 VERSION
 KEYWORDS
 SOURCE
 ORGANISM
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Archosauria; Aves; Neognathae; Galliformes; Phasianidae; Phasianinae; Gallus.
 1 (bases 1 to 2962)
 Guan,W. and Condit,M.L.
 Characterization of Netrin-1, Neogenin and cUNC-5H3 expression during chick dorsal root ganglia development
 Gene Expr. Patterns 3, 369-373 (2003)
 2 (bases 1 to 2962)
 Guan,W. and Condit,M.L.
 Direct Submission
 Submitted (26-NOV-2002) Neurobiology & Anatomy, University of Utah, 20 North, 1900 East, Salt Lake City, UT 84132-3401, USA

FEATURES
 location/Qualifiers
 1..2962
 /organism="Gallus gallus"
 /mol_type="mRNA"
 /db_xref="taxon:9031"
 1..2796
 /codon_start=1
 /product="UNC5-like protein 3"
 /protein_id="AA067275.1"
 /db_xref="GI:31442351"
 /translation="MGKLEGTAAACGLGMGYLLHVVLPALAVLGASRPGSAQQDD
 PFHELPEFPDPEPELPHLEPEEATVIVKKNVNLKASPAATQIYFKCNSEWVHQ
 KDVVDREVTSLGIVCEVSEIHSRQVEELFGPEYDWCQCVAWSSAGTTSKRYAY
 RIAYLRKTFEQLGKEYSLQEVLLQCRPEGPVAVELWLNKEVIDPVDNRFYI
 TIDHNLITKQARLSDTANYTCVAKNIVAKRSTTATVIVVNGWSTWESACNRSR
 GRGFKRTRTCTNPAPLNGAFCEGNOVKIATCTLCPVDGKWSWSKSTGCTGCTH
 WRRETPAPKNGSKOCEGLVLOKNCITGLCQAAAPSDSDVALYGVIVAVICLA
 ISVVVAFVYKRNHRDESIIIDSLNGFPVNIKAARQDLAVPDLTSAAMYR
 GPVALHDVSDKIPTMNSPILDPNFKIKYNTSGAVTPQDELSDFSKLSQITQS
 LLENETLNKNSLARQDPTCTAFNGLSGHLVINSVGLIIPAGVPOGRVYE
 MYVTVHREKMEPPVEDSOTLLTPVSCPPCALLTRPVLTTHHCAENMDQIOL
 KHAQGGWEDVVVWRENTPCYIOLDPEACHLTETLSTYALVGSQITKAAAKRL
 KLAIFPLSCSSLSYRVCLDQDALKVLQIROMGGOLLEPFLHFKGSTHN
 LRLSHIDLAHSLWKSUKLPKATQEIPIFYHWSGQNRNHCFTLRFSLNTLFLKCLC
 VRQVEGGQIFQLNCSVEEPTGDIYFIMDSAGSIITVGNFASIPILPIQKLCSSL
 DAPQTRGHDNRMLAHKLDRLYLNYFATKSPTSPTGVILDLWEAQNPDPGNLSMLA
 VMGRHETVTVSLAAEGNY"

ORIGIN
 Query Match 32.6%; Score 938.4; DB 5; Length 2962;
 Best Local Similarity 61.3%; Pred. No. 2.2e-145;
 Matches 1637; Conservative 0; Mismatches 961; Indels 72; Gaps 5;

QY 190 TGCCTGTGCGCAACCCGAGCCTGCTTCCCACTTCTCGTGGTGGAGCCGAGATGTGTACA 249
 Db 158 TTCTCTGATCTCTCCAGGCCATTTGCCCATCTTCTCAITGAACCCGAGAGCTTACA 217

QY 250 TCGTCAAGAAACAGCCAGTGTCTGTGTGCGAAGCGCGTCCCGCCACGACGATCTTCT 309
Db 218 TCGTGAAGAAACAGCCAGTGTCTGTGTGCGAAGCGCGTCCCGCCACGACGATCTTCT 277
QY 310 TCAAGTGCACGCGGAGTGGGTGCGCCAGGTGCGACAGTGTGATCAGCGCAGCAGACG 369
Db 278 TTAAGTGCACAGAGTGAATGGGTTCATCAGAAAGATCATGTGTGATGAGAGTAGATG 337
QY 370 GGAGCAGTGGTGCAGCCGACCATGAGAGTCCGCAITTAATGTCTCAAGGCGACGAGTGCAGA 429
Db 338 AAACCTCTGTCTGTCTGTGAGGTGAGCATCGAGATTTCCCGCAGCAGGTGGAAG 397
QY 430 AGGTGTCGGGTGGAGGAATACTGGTGCAGTGGGTGCGATGAGACTCTCCGGACCA 489
Db 398 AGCTCTTTGGACCCGAGGACTACTGGTGCAGTGTGTCCCTGGAGCTCAGCTGGACCA 457
QY 490 CCAAGAGTCAGAAAGCCTACATCCGATAGCCAGATTGGCGCAAGAACTTCAGACGAGGAGC 549
Db 458 CCAAGAGCCGCAAGGCTAGTCCGCAITTGCAITATCTCAGAAAGACTTTTGACGAGGAGC 517
QY 550 CGCTGCGCAAGGAGTGTCCCTGGAGCAGGGCATCGTGTGCCCTGCGCTCCACCGAGG 609
Db 518 CGCTGGGAAAGAGTGTCCCTGGAGCAAGAGTCTCTGCTCCAGTGCCTCTCTGAAG 577
QY 610 GCATCCCTCCAGCGGAGGTGGAGTGGCTCCGGAAACAGAGACCTGGTGGACCCGCTCG 669
Db 578 GCATTCAGCTAGCTGAGGTAGAGTGGCTGAAGAAATGAAGAGGTGATCGATCTCTGGAAG 637
QY 670 ACCCAATCTATACATACCGCGGAGCACAGCTGGTGGCGACAGGCGCGCTTGCTG 729
Db 638 ACCGAATTTTACATCACCATTTGATCAACCTGATCATCAAGCAAGCCCGGCTTCCG 697
QY 730 ACACGGCCAACTACACTGCTGGGCCAAGAAACATCGTGGCAGCTGCGCGAGGCGCTCG 789
Db 698 ACACGGCTAACTACACTGTGTGGCCAAACATTTGGGCCAAAGAAAGAACACACAG 757
QY 790 CTGCTGTATCTGTCTAGTGAACGGTGGGTGGTGGAGCTGGACCGAGTGGCTGCTGCA 849
Db 758 CAACTGTGATGTCTATGTGAATGGAGGCTGGTCTTACCTGGACCGAGTGGTCAAGCTGCA 817
QY 850 GCGCCAGCTGTGGCGCGCTGCGAGAAACGAGCGGAGCTGCACCAACCCCGCGCTC 909
Db 818 ACAGCGCTGTGGAGAGGCTTCCAGAGCGCACAGGACCTGCATNACCTTGCCTCCAC 877
QY 910 TCAACGGGGCGCTTTCTGTGAGGGCGAGAAATGTCATGACCGCAACGCTCTCTCTGTC 969
Db 878 TCAATGGGGGGCTTCTCTGAGGGGCAAAATGTTCAAGAAATAGCTTGCACCACTGT 937
QY 970 TTGTCTCTGAGACGGAGCTGGAGCCGTTGGAGCAAGTGGTGGCTGGGCTGCACT 1029
Db 938 GT---CCAGTGGATGGCAAAATGACGCTCTGGAGCAAGTGGTCCACTTGTGGCACAGAT 994
QY 1030 GCACCCACTGGCGAGCGGTGAGTGTCTGCACCCAGCACCCCGCAACGAGGGGAGGAT 1089
Db 995 GTACCCACTGGCGGAGGAGTGCACAGCTCGGCCCCGAAGATGGAGGCAAGGACT 1054
QY 1090 GCCAGGSCACTGACCTGGACCCCGCAACTGTATACAGTACCTCTGTGTATACAGTGTCT 1149
Db 1055 GTGAGGAGCTGGTGTCTGACGTCTAAGAACTGCATGTATGGGCTCTGCATGACGCTGCAC 1114
QY 1150 CTGGCCCTGAGGAGCTGGCCCTCTATGTGG---CCTCATCGCGTGGCGCTGCGCTGG 1206
Db 1115 CTGACTCGGATGATGTGTCTCTACGTGGGATGTGATGTGTGTGTGTGTGTGTGTGTGT 1174
QY 1207 TCCTGCTGCTGTCTCTCATCTCTCTGTTTATTCGGGAAAGAGGAGGCGGTGCACTCAG 1266
Db 1175 CTAATTTCTGT 1234
QY 1267 ATGTGGGTGACTCGTTCATCTACCTCAGGCTTCAGGCGGCTCAGCATCAAGCCAGCA 1326
Db 1235 ATATTATCGATCATCGCGCTAAATGGGGATTTCAAGCTGTGATCAATCAAGGCTGCA 1294

QY 1327 AAGCAGACAAACCCCATCTGTCTCACCATCCAGCCGAGCTCAGCACCAACCACTACC 1386
Db 1295 GACAGAC-----CTCTGGCAGTGCACACAGACTCCTCCTCTCTCTGTCAGCATGT 1345
QY 1387 AGGGCAGTCTCTGTCCCGGAGGATGGGCCAGCCCAAGTTCAGCTCACCATGGGC 1446
Db 1346 ACAGGGGCGCTGTGTATGCTTGTGATGTCTCTGTATAAATCCCAATGACCAATCTC 1405
QY 1447 ACCTCTCAGCCCTCGGTGGGCGGCCACACACTGCACACAGCTCTCCCACTCTG 1506
Db 1406 CGATCTCGACCCCACTGCCCCAATCTGAAGATTAAAGTTTATACACCTCTGAGCAGTCA 1465
QY 1507 AGGCGGAGAGTTCGTCTCTCCGCTCTCCAC----- 1538
Db 1466 CCCCCCAGGATGAATCTCTGACTTCTCTCCTCAAGCTGTCCCCACAGATTACCCAGTCTC 1525
QY 1539 -----CAGAACTACTTCGGCTCCCTGCCCCCGGAGGACCA 1572
Db 1526 TGTGGAGAAATGAGACTCTGAAACGTGAAGAACCAAGCCTTGACGGCAACAGACCCAT 1585
QY 1573 GCAACATGACCTATGGACCTTCAACTTCTCGGGGCGCGGTGATGATCCCTAATACAG 1632
Db 1586 CTTGCACTGCAATTTGGACCTTCACTCGTTAGGGGCGCACCTAGTAAATCTTAATTCAG 1645
QY 1633 GTATCAGCTCTCTCATCCCCCAGATGCCATACCCCGAGGAGATCTATGAGATCTACC 1692
Db 1646 GAGTGAAGCTTGTGATCCAGCAGGGCTGTTTCCCCAAGGAAGATCTATGAAATGTATG 1705
QY 1693 TCACGCTGCACAAACCGGAGAGAGTGGTGGCTTGGCTTGTGAGCTGTGAGCCCTCTGA 1752
Db 1706 TGACAGTCCACAGGAAGGAGGAGCATGAGACCACTGTAGAACACAGCCAGACCTGCTGA 1765
QY 1753 GTCCCATCTGTAGCTGTGACCCCTGGCGTCTGTCTCACCCGCGCAGTCACTCTGGCTA 1812
Db 1766 CACCACTGTGAGCTGTGCGCCACAGAGCGCTGTGACCCGACCGTGTGCTGACCA 1825
QY 1813 TGACACACTGTGGGAGCCCGCTGACAGCTGGAGCTGGCCCTCAAAAAGAGTGTCTG 1872
Db 1826 TGCACCACTGTGCTGAGCCCAACATGGATGATGGCAGATCCAGCTCAAGCACAGGACG 1885
QY 1873 GCGAGGAGCTGGGAGCAGGATGTGCTGACCTGGCGAGGAGGCGCCCTCCCACTCT 1932
Db 1886 GCGAGGAGCATGGGAGGATGTAGTGGTGGTCT---GGGAGGAAACCTTCAACCACTCCAT 1942
QY 1933 ACTACTGCCAGCTGGAGGCGAGTGTGCTGTCTTACCGAGCAGCTGGGCGGCTTTG 1992
Db 1943 GCTATATCCAGCTGGACCCAGAGGCTGTCTATATCTGACGAGACCTCAGCACGTAAG 2002
QY 1993 CCTGTGTGGAGAGGCGCTCAGCGTGGTGGCGCCAAAGCGCTCAAGCTGCTTCTGTTG 2052
Db 2003 CCTGTGTGGACAATCCATCACAAAGCAGCAGCCCAAAAGCTCTCAAAATGGCCATCTTTG 2062
QY 2053 CGCCGCTGGCTGCACCTCCCTGAGTACAAATCGGGTCTACTGCTGCATGACACAC 2112
Db 2063 GACCACTGTCTCTCTCACTGGATACAGCATCCGCGTCTACTGCTCGATGACACAC 2122
QY 2113 ACATGCACTCAAGGAGGTGGTGCAGCTGGAGAGCAGCTGGGGGAGCAGCTGATCCAGG 2172
Db 2123 AGATGCCCTGAAGAGGTCTCTCAGCTTGGCGGAGATGGGTGGGAGCTGTGTGAGG 2182
QY 2173 AGCCACGGGTCTGCACTTCAAGGACGATTACACAACTGGCGCTTATCCATCCACGATG 2232
Db 2183 AACCCAAAATTTTGCATTTTAAAGGAAGTACCCCAACCTGCGCTTATCCATTCATGACA 2242
QY 2233 TGCCGAGCTCCCTGTGGAGAGTAAAGCTCTTGTGCTAGCTACCAAGGAGATCCCTTTTATC 2292
Db 2243 TTGCCCACTCTCTGTGAAGAGCAAACTGCGGCTAAATACAGGAGATTTCTTTTATACC 2302
QY 2293 ACATCTGGAATGGCAGCAGCGGTACTTGCATGCACTTGCACCTTACCTGGAGGCTGTGAGCC 2352
Db 2303 ACATCTGGAGTGGTGCAGAGGAACCTTGCACTGCACTTTCACGCTGGAACGATTCAGTC 2362
QY 2353 CCAGCACTAGTGACCTGGCCTGTCAAGCTGTGGGTGTGGAGGTCGAGGCGGACGCGGAGA 2412

```

Db      2363  TCATACCTCTGGAGCTGCTGTCGAAACTCTGTGTGCGCAAGTCGAAGGAGAGGGCAGA 2422
Qy      2413  GCTTCAGCATCAACTTCAACATCACCAGGACACAAAGTTTGTGAGCTGCTGCTGG 2472
Db      2423  TCTTCAGCTGAATGCTCAGTATCAGAGGAACCACTGGCATTTGATTTATCCCATATGG 2482
Qy      2473  AGAGTGAAGCGGGTCCCAGCCCTGTGTGGGCCCGAGTGCCTTCAAGATCCCTTCTCTCA 2532
Db      2483  ATTTCAGCAGCAGCATCACAAAGATAGTTGGGCCCAAGGTTTCAGCATCCCTTCCAA 2542
Qy      2533  TTCCGAGAGAGATAATTTCAGCTCTGGAACCCACCTGTAGCGGGGTGCGACTGGCGGA 2592
Db      2543  TAAGCAGAGAGCTCTGAGCAGGCTGTGATGACCCAGAGCCCGGGGCCATGACTGGAGGA 2602
Qy      2593  CTCTGCCCCAGAAACTCCACTGAGAGCCATCTCAGCTTCTTGTGCTTCAAGCCAGCC 2652
Db      2603  TGTGCCCCACAGCTGAATTGGACAGTACCTAAATTAATTTGTACGAAGTCGAGTC 2662
Qy      2653  CCACAGCCATGATCTCAACTGTGGGAGCGGCGACTTCCCAACGGCAACTCAGCC 2712
Db      2663  CCACTGGGGTGTCTCTGGATCTCTGGAGGCCAGATTTCCCTGATGGCAACTGAGCA 2722
Qy      2713  AGCTGCTGCAGCAGTGGTGGTGGTGGCCAGCAGCAGCTGGCTCTTTCACAGTGTGG 2772
Db      2723  TGTGTCGACAGCTGTGGAAGAAATGGAGCAGATGAAACCGTGTCTTCTTGGCAGCAG 2782
Qy      2773  AGCTGAGTGTCTGAGCGCGCCAGGCGCCGA 2802
Db      2783  AAGGAATTAATGATGACAGCTTGGCGTGA 2812

```

RESULT 15

```

LOCUS      AB118026                      9328 bp    mRNA    linear    ROD 19-DEC-2003
DEFINITION Rattus norvegicus Unc5h3 mRNA, complete cds.
ACCESSION  AB118026
VERSION    AB118026.1  GI:40217509
KEYWORDS
SOURCE     Rattus norvegicus (Norway rat)
ORGANISM  Rattus norvegicus

```

```

REFERENCE  1
AUTHORS   Kuramoto,T., Kuwamura,M. and Serikawa,T.
TITLE     Rat neurological mutations cerebellar vermis defect and hobble are caused by mutations in the netrin-1 receptor gene Unc5h3
JOURNAL   Mol. Brain Res. (2003) In press

```

```

REFERENCE  2  (bases 1 to 9328)
AUTHORS   Kuramoto,T. and Serikawa,T.
TITLE     Direct Submission
JOURNAL   Submitted (22-AUG-2003) Takashi Kuramoto, Institute of Laboratory Animals, Graduate School of Medicine, Kyoto University; Yoshidakone-cho, Sakyo-ku, Kyoto 606-8501, Japan
          (E-mail:tkuramoto@anim.med.kyoto-u.ac.jp,
          URL:www.anim.med.kyoto-u.ac.jp, Tel:81-75-753-4494,
          Fax:81-75-753-4409)

```

FEATURES

```

source    1..9328
           /organism="Rattus norvegicus"
           /mol_type="mRNA"
           /db_xref="taxon:10116"
           /chromosome="2"
gene      1..9328
           /gene="Unc5h3"
           /protein_id="BAD05181.1"
           /codon_start=1
           /db_xref="GI:40217510"
           /translation="MRKGLRATAARCGLGLYLLQMLVLPALALLSASGTGSAQQDDDD
           FFHELPTFSDPPEPLPHLEPEAYIVKNKPNLYCKASPATQIYFKCNSEWHQ

```

```

3'UTR      2947..9328
           /gene="Unc5h3"
           /note="simple sequence repeats"
           /rpt_unit="gt"
           /polya_signal
           /gene="Unc5h3"

```

ORIGIN

```

Query Match      32.5%; Score 936.8; DB 10; Length 9328;
Best Local Similarity 61.8%; Pred. No. 3.6e-145;
Matches 1631; Conservative 0; Mismatches 937; Indels 72; Gaps 6;

QY      204  CCGGACCTGCTTCCCACTTCTCTGGTGGAGCCGAGGAGTGTACATCTCTCAAGTGCACCGG 323
Db      322  CTTGAGCCATTACCACTTCTCTATTGAGCCGAGGAGTGTACATCTCTCAAGTGCACAG 381
QY      264  CCAGTGTCTTGTGTGCAAGCCGTCGCCGACGACAGATCTTCTTCAAGTGCACCGG 323
Db      382  CCTGTGAACCTGATTGTAAGCCAGCCCTGCCACCCAGATCTACTTCAAGTGCACAGC 441
QY      324  GAGTGGTGGCGCCAGTGGACCACTGATCTGAGCCGACACAGACGGGAGCAGTGTGAG 383
Db      442  GAATGGGTTTCATCAGAAGGACCACTGATGACGAGAGGGTAGATGAACCTCCGGTCTC 501
QY      384  CCGACCATGGAGTCCGCTTAATGCTCTCAAGGACGAGTGTGAGAGTGTTCGGGCTG 443
Db      502  ATTGTGAGAGAGTGAAGTATGAGATTTACCCAGCAGGTGGAGAACTGTTCCGGGCT 561
QY      444  GAGGAATCTGTGTCAGTGGCTGTCATGGAGCTCTCTGGGACCAACCAAGAGTCAAG 503
Db      562  GAAGATTACTGTGTCAGTGTGGCTGGAGTCTCAGCGGCACTACGAGAGTGGAG 621
QY      504  GCCTACATCCGATAGCCAGATTGGCGAAGAACTTCGAGCAGGAGCCGCTGGCCAGGAG 563
Db      622  GCATACGTGGCATTTGCGTATCTCGGAAGACATTTGAGCAGGAACCCCTTGGGAAAGAA 681
QY      564  GTGTCCCTGGAGAGGATCGTGTGCTGCTGCCCTGCCACCGAGGGGATCCCTCCAGCC 623
Db      682  GTGTCTTGGAGCAGGAGTGTACTCAGTCCCGCCCACTTGAAGGATCCCAATGGCT 741
QY      624  GAGTGGAGTGGCTCCGGAACGAGACCTGTTGGACCCGCTCCCTGGAGCCCAATGTATAC 683
Db      742  GAGTGGAGTGGCTAAAGAATGAAGACATAATTGATCTGTGAGAGATCGGATTTTAT 801
QY      684  ATCAGCGGAGCAGACCTGTTGGTGGCAGACGCGCCGCTTGTGACACGCGCACTAC 743
Db      802  ATCACTATCGATCAACCTGATCATCAAGCAGGACGACACTCTCAGATACAGCAATAT 861
QY      744  ACCTGCGTGGCCAAAGAACATCGTGGCAGCTCCGCGCAGCGGCTCCGCTGTCTCGTC 803
Db      862  ACCTGTGTGCCAAAGAACATTTGTCAGAGGAAGAACCAACAGCCACCGTCATCGTG 921
QY      804  TACGTGAACCGTGGTGGTGGACGTGGAGCGAGTGGTCCGCTCTCAGCGCCAGCTGTGG 863
Db      922  TATGTTAATGTGGTGGTCCACCTGGGAGAAATGGTCTGTGTGTAAACGCGCTGTGG 981
QY      864  CGCGGCTGGCAGAAACGAGCCGAGCTGCACCAACCCGGCGGCTCTCTAAACGGGGCGCT 923

```

```
Db 982 CGAGGATATCAGAAACGACAGAACCTGACCAACCCAGCCCTCAATGGTGGGCC 1041
QY 924 TTCTGTGAGGGGAGAAATGTCATACCGCACCGTCTCTCTCTGTGTCTCTGTGGAC 983
Db 1042 TTCTGCGAGGGGAGAGGTGTACAGAA--AATAGCATGCACACGTATGCCAGTGGAT 1098
QY 984 GGCAGCTGGAGCCCTGGAGCAAGTGGTGGGCTGTGGGCTGGACTGCACCCACTGGGG 1043
Db 1099 GGCAGGTGGACTTCTGTGAGCAAAATGGTCAACTGTGGGACTGAGTGACCCACTGGCG 1158
QY 1044 AGCCGTGAGTGTCTGACCCAGCACCCGACAGGAGGGAGAGTGCAGGCGCACTGAC 1103
Db 1159 AGGAGGGAGTGTACAGCACCGCCCAAGAACCGGGGTAAAGACTGTGATGCCGTGTC 1218
QY 1104 CTGGACACCCGCAACTGTATACCACTGACCTCTGTGTACACAGTGTCTTCTGGCCCTGAGGAC 1163
Db 1219 CTCCATCCAGAACTGACCGATGGCTGTGCATGCAGGCTGCTCCTGACTCAGATGAC 1278
QY 1164 GTGGCCCTCTATGTGGG---CCTCATCGCGGTGGCCGTCTGCTGTGCTGTGCTGT 1220
Db 1279 GTTGTCTCTACGTTGGGATTTGTATCGCTGTAAACCGTTTGTCTGGCAATCACTGTGTG 1338
QY 1221 GTCTCTATCTCTGTTTATTTGCGGAGAGAGGGGCTGACTCAGATGTGGCTGACTCG 1280
Db 1339 GTGGCCCTGTCTGTGTATCGCAAGAACCCCGTGACTTTGAGTCTAATATCATGACTTC 1398
QY 1281 TCCATTTCTCACCTCAGGCTTCCAGCCCGTTCAGCATCAAGCCCAAGCAAGACAGAACCC 1340
Db 1399 TCAGCACTCAATGGCGCTTTCAGCTGTGAACATCAAG-----GCTGCCAGACAA 1449
QY 1341 CATCTGCTCACATCCAGCCGGACCTTCAGCACACACACACCTACAGGGGAGTCTCTGT 1400
Db 1450 GATCTCTGCGAGTACCCCGGATCTCACCTCAGCTGCGAGCCATGTACAGGGGACCCGTC 1509
QY 1401 CCCCAGGAGTGGGCCCGAGCCCAAGTTCAGCTCACCAATGGGCACTGTCTAGCCCC 1460
Db 1510 TATGCTCTCATATGTCTCAGACAAATCCCAATGACCAATCTCAATCTAGACCCA 1569
QY 1461 CTGGTGGCGGCGGCACACACTGCACACAGCT-----CTCCCACTCTGAG 1508
Db 1570 CTGCCCAATTTGAATCAAGGTGTACACAGCTCAGGTGCTGTCACTCTCAGGACGAC 1629
QY 1509 GCGGAGAGTGTCTCCCGCTCTCCACCGAGACTACTTCCGCTCCCTGCCCCGAGGC 1568
Db 1630 CTGTGCTGAGTCTCATCCAAGCTGTACCCCGAGATGACCAATCTTGTGTAGAGATGAG 1689
QY 1569 ACCAGCAACATGA-----CCTAT 1586
Db 1690 GCCTGAACTGAGAACAGAGTCTCGAAGACAGACTGACCCATCCTGCAAGCGTTT 1749
QY 1587 GGGACCTTCAACTTCTCGGGGCGGCTGATGATCCCTAATACAGGTATCAGCTCCTC 1646
Db 1750 GGCACCTTCAACTCCCTGGCGGTCACTCATCATCCCTAATCAGGAGTAACTGTCTG 1809
QY 1647 ATCCCCCAGATGCCATACCCGAGGAGAGATCTATGAGATCTACCTCAGCTGCACAG 1706
Db 1810 ATTCCCGCTGGGCGCATCTCTCAGGGAGAGTCTATGAATGTATGTGCTGTACACAG 1869
QY 1707 CCGGAGAGCTGAGGTTGCCCTAGCTGGCTGTGAGACCCCTGCTGAGTCCCATCGTTAGC 1766
Db 1870 AAGAGAAATATGAGCCCCCATGGAAGATCTCAGACCCCTGCTGACCCCTGTGTGAGC 1929
QY 1767 TGTGGACCCCTGGCGTCTGCTCTCACCGGCGAGTCAATCTGGTATGGACCACTGTGGG 1826
Db 1930 TGTGGGCTCTCTGGAGCTCTGTTGACCGCTCTGTCTATCTCTCACTCTGCTACTGTGCA 1989
QY 1827 GAGCCAGCCCTGACAGTGTGAGCTGCGCTCAAAAGAGAGTGTGCGAGGGGAGCTGG 1886
Db 1990 GACCTTAACACCGAGGACTGGAAGANTCCAGCTCAAAACACGAGGTGCGAGGACAGTGG 2049
QY 1887 GAGCAGGATGTGTGCTGACCTCGGGGAGGAGGCGCCCTCCACCTCTACTACTGCGAGCTG 1946
```

Search completed: October 5, 2004, 18:44:04
Job time : 11157 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: October 5, 2004, 13:21:05 ; Search time 1076 Seconds

(without alignments)
11374.590 Million cell updates/sec

Title: US-09-970-944-1

Perfect score: 2881

Sequence: 1 agctggggtccggctgtag.....ccttccccacacgggggaga 2881

Scoring table: IDENTITY NUC

Gapop 10.0 , Gapext 1.0

Searched: 3373863 seqs, 2124099041 residues

Total number of hits satisfying chosen parameters: 6747726

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : N_Geneseq_29Jan04.*

1: Geneseqn1980s.*

2: Geneseqn1990s.*

3: Geneseqn2000s.*

4: Geneseqn2001as.*

5: Geneseqn2001bs.*

6: Geneseqn2002s.*

7: Geneseqn2003as.*

8: Geneseqn2003bs.*

9: Geneseqn2003cs.*

10: Geneseqn2004s.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	2881	100.0	2881	6	ABK49422
2	2676.4	92.9	2752	6	ABK37922
3	2615	90.8	2697	6	ABK52891
4	2402.6	83.4	3580	6	ABK15169
5	2296.2	79.7	2907	4	AAK52261
6	2235.8	77.6	3014	2	AAV52940
7	2189.4	76.0	2697	6	AAK16843
8	1570.4	54.5	1787	2	AAV52941
9	1291.2	44.8	9700	4	AAK90958
10	1193.6	41.4	1321	4	AAH99530
11	916	31.8	3646	5	AAH57538
12	889	30.9	2860	6	ABT06279
13	887.4	30.8	2860	6	ABT06280
14	875.8	30.4	2995	6	ABK92062
15	866.4	30.1	2895	6	AAQ93898
16	853.6	29.6	3884	4	AAQ32116
17	853.6	29.6	3884	7	ACD23925
18	853.6	29.6	3884	7	ACA67066
19	853.6	29.6	3884	7	ACA03675
20	853.6	29.6	3884	7	ABX89213
21	853.6	29.6	3884	7	ACD41867
22	853.6	29.6	3884	7	ACA04096
23	853.6	29.6	3884	8	ADA45664

24	853.6	29.6	3884	8	ADA76095	Ada76095	Human	PRO
25	853.6	29.6	3884	8	ADA18745	Ada18745	Human	PRO
26	853.6	29.6	3884	8	ADA61368	Ada61368	Homo sapi	
27	853.6	29.6	3884	8	ADB19153	ADB19153	Novel	hum
28	853.6	29.6	3884	8	ADB27694	ADB27694	Novel	hum
29	853.6	29.6	3884	8	ADA86173	ADA86173	Novel	hum
30	853.6	29.6	3884	8	ADB15737	ADB15737	Novel	hum
31	853.6	29.6	3884	8	ADA47523	ADA47523	Human	PRO
32	853.6	29.6	3884	8	ADA67318	ADA67318	Human	PRO
33	853.6	29.6	3884	8	ADB30325	ADB30325	Novel	hum
34	853.6	29.6	3884	8	ADA85621	ADA85621	Novel	hum
35	853.6	29.6	3884	8	ADA96833	ADA96833	Human	PRO
36	853.6	29.6	3884	8	ADA79137	ADA79137	Human	PRO
37	853.6	29.6	3884	8	ADA87276	ADA87276	Novel	hum
38	853.6	29.6	3884	8	ADA91570	ADA91570	Novel	hum
39	853.6	29.6	3884	8	ADB16478	ADB16478	Human	PRO
40	853.6	29.6	3884	8	ADB14633	ADB14633	Human	PRO
41	853.6	29.6	3884	8	ADB18594	ADB18594	Novel	hum
42	853.6	29.6	3884	8	ADA93809	ADA93809	Human	PRO
43	853.6	29.6	3884	8	ADB19705	ADB19705	Novel	hum
44	853.6	29.6	3884	8	ADB13017	ADB13017	Human	PRO
45	853.6	29.6	3884	8	ACD98496	ACD98496	Novel	hum

ALIGNMENTS

RESULT 1

ABK49422

ID ABK49422 standard; DNA; 2881 BP.

XX AC ABK49422;

XX DT 15-JUL-2002 (first entry)

XX DE DNA encoding human UNC5-like protein NOV1.

XX KW Human; NOVX polypeptide; cardiomyopathy; atherosclerosis; cancer;

XX KW cell signal processing; metabolic pathway modulation; cancerous tissue;

XX KW antibody; diabetes; transgenic animal; UNC5-like protein; NOV1;

XX KW chromosome 13; gene; ds.

XX OS Homo sapiens.

XX FH Key

XX FT CDS

XX FT Location/Qualifiers

XX FT 87..2786

XX FT /*tag= a

XX FT /product= "Human UNC5-like protein NOV1"

XX WO200229038-A2.

XX PD 11-APR-2002.

XX PF 04-OCT-2001; 2001WO-US031377.

XX PR 04-OCT-2000; 2000US-0237862P.

XX PA (CURA-) CURAGEN CORP.

XX PI Herrmann JL, Rastelli L, Shimkets RA;

XX DR WPI; 2002-340104/37.

XX DR P-ESDB; AAU79939.

XX PT Novel isolated NOVX polypeptide, and encoded polynucleotide, useful for

XX PT treating cardiomyopathy, atherosclerosis, and cancer.

XX PS Claim 8; Page 7-8; 180pp; English.

XX CC The present invention relates to a new NOVX polypeptide having a 900

XX CC (NOV1), 4349 (NOV2), 940 (NOV3), 798 (NOV4), 965 (NOV5), or 331 (NOV6)

XX CC residue amino acid sequence, as given in the specification. The novel

XX CC polypeptide, and its encoding polynucleotide, are used to treat

CC cardiomyopathy, atherosclerosis, cancer or a disease related to cell
CC signal processing and metabolic pathway modulation, in a human. Detecting
CC the polypeptide or polynucleotide is useful for identifying cancerous
CC tissue. The antibody can be used to treat diabetes or cancer. The host
CC cells can be used to produce non-human transgenic animals useful in drug
CC screening. The present nucleic acid sequence is that of the human UNC5-
CC like NOV1 gene located on chromosome 13. This sequence encodes the human
XX UNC5-like protein NOV1 of the invention

SQ Sequence 2881 BP; 526 A; 985 C; 868 G; 502 T; 0 U; 0 Other;

Query Match 100.0%; Score 2881; DB 6; Length 2881;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 2881; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 AGCTGGGGCTCCGGGCTGAGGCGCTAAAGCCGCCCTCCCGCCCGGGGGCCCCGGGCGCG 60
DB 1 AGCTGGGGCTCCGGGCTGAGGCGCTAAAGCCGCCCTCCCGCCCGGGGGCCCCGGGCGCG 60

QY 61 GCCCGCCCGCTCCCGCCCGGGCCATGCGCGTCCGGCCCGGCTGTGGCCAGCGCTCC 120
DB 61 GCCCGCCCGCTCCCGCCCGGGCCATGCGCGTCCGGCCCGGCTGTGGCCAGCGCTCC 120

QY 121 TGGGCATAGTCTCGCGCTTGCTCCGGCTCGGGTCCCGAGAGTGCACCGTGG 180
DB 121 TGGGCATAGTCTCGCGCTTGCTCCGGCTCGGGTCCCGAGAGTGCACCGTGG 180

QY 181 CCAACCCAGTGTGGTGCACCCGGACCTGCTCCCGACCTTCCCTGGTGGAGCCCGAGG 240
DB 181 CCAACCCAGTGTGGTGCACCCGGACCTGCTCCCGACCTTCCCTGGTGGAGCCCGAGG 240

QY 241 ATGTGTACATGTCGAAGAACAGCCAGTGTGTTGTGTGCAAGGCCGTGCCCGCACGC 300
DB 241 ATGTGTACATGTCGAAGAACAGCCAGTGTGTTGTGTGCAAGGCCGTGCCCGCACGC 300

QY 301 AGATCTTCTCAAGTCAACGGGAGTGGTGGCCAGGTGGACACAGTGCATCGAGCGCA 360
DB 301 AGATCTTCTCAAGTCAACGGGAGTGGTGGCCAGGTGGACACAGTGCATCGAGCGCA 360

QY 361 GCACAGCGGGAGCAGTGTGAGCCGACCATGAGGTCCGCAATTAATGTCTCAAGCAGC 420
DB 361 GCACAGCGGGAGCAGTGTGAGCCGACCATGAGGTCCGCAATTAATGTCTCAAGCAGC 420

QY 421 AGTTCAGAGAGTGTTCGGGCTGGAGGAATACGTGTGCCAGTGGCATGAGCTCCT 480
DB 421 AGTTCAGAGAGTGTTCGGGCTGGAGGAATACGTGTGCCAGTGGCATGAGCTCCT 480

QY 481 CGGGCCACCACCAAGAGTCAAGAGCCCTACATCCGCATAGCCAGATTGCGCAAGAACTTCG 540
DB 481 CGGGCCACCACCAAGAGTCAAGAGCCCTACATCCGCATAGCCAGATTGCGCAAGAACTTCG 540

QY 541 AGCAGGAGCGCTGGCCAAAGAGGTGTCCCTGGAGCAGGSCATCGTGTGCCCTGCGGTC 600
DB 541 AGCAGGAGCGCTGGCCAAAGAGGTGTCCCTGGAGCAGGSCATCGTGTGCCCTGCGGTC 600

QY 601 CACCGAGGCGATCCCTCCAGCCGAGTGGAGTGGCTCCGGAACAGGAGCTGTGSGACC 660
DB 601 CACCGAGGCGATCCCTCCAGCCGAGTGGAGTGGCTCCGGAACAGGAGCTGTGSGACC 660

QY 661 CGTCCCTGGACCCCAATGTATACATACGCGGGAGCACAGCTGTGGTCCGACAGGCC 720
DB 661 CGTCCCTGGACCCCAATGTATACATACGCGGGAGCACAGCTGTGGTCCGACAGGCC 720

QY 721 GCCTTGTGACACGGGCCAACTACCTGTGCGCAAGAACATCGTGGGACGTGCGCGCA 780
DB 721 GCCTTGTGACACGGGCCAACTACCTGTGCGCAAGAACATCGTGGGACGTGCGCGCA 780

QY 781 GCGCCTCCGCTGTGTCTCATGTGTAACCGTGGTGTGCACTGAGCAGCGAGTGT 840
DB 781 GCGCCTCCGCTGTGTCTCATGTGTAACCGTGGTGTGCACTGAGCAGCGAGTGT 840

QY 841 CCGTCTGCAGCGCCAGCTGTGGGCGCGCTGGCAGAACGAGCCGGAGCTGCACCAACC 900
DB 841 CCGTCTGCAGCGCCAGCTGTGGGCGCGCTGGCAGAACGAGCCGGAGCTGCACCAACC 900

DB 841 CCGTCTGCAGCGCCAGCTGTGGGCGCGCTGGCAGAACGAGCCGGAGCTGCACCAACC 900

QY 901 CGGCGCCTCTCAACCGGGGGCGCTTCTGTGAGGGGCGAGAATGCTCCATGACCGCAGCTCT 960
DB 901 CGGCGCCTCTCTCAACCGGGGGCGCTTCTGTGAGGGGCGAGAATGCTCCATGACCGCAGCTCT 960

QY 961 CCTCTCTCTTGTCTCTGTGACGCGCAGCTGAGCGCCGTGGAGCAAGTGTGCGGCTGTG 1020
DB 961 CCTCTCTCTTGTCTCTGTGACGCGCAGCTGAGCGCCGTGGAGCAAGTGTGCGGCTGTG 1020

QY 1021 GGTGTGACTGCACCCACTGGGAGCGCGTGTGTGTGTGACCCAGGACCCCGCAACGAG 1080
DB 1021 GGTGTGACTGCACCCACTGGGAGCGCGTGTGTGTGTGACCCAGGACCCCGCAACGAG 1080

QY 1081 GGGAGAGTGCACGGGCACTGACCTGGACACCCGCACTGTACCACTGACCTGTGTGTAC 1140
DB 1081 GGGAGAGTGCACGGGCACTGACCTGGAGACCCCGCACTGTACCACTGACCTGTGTGTAC 1140

QY 1141 ACAGTGTCTTCTGGCCCTGAGGACGTGGGCCCTCTATGTGGGCCCTCATGCGCGTGGCGTCT 1200
DB 1141 ACAGTGTCTTCTGGCCCTGAGGACGTGGGCCCTCTATGTGGGCCCTCATGCGCGTGGCGTCT 1200

QY 1201 GCTGTGCTGT 1260
DB 1201 GCTGTGCTGT 1260

QY 1261 ACTCAGATGTGGCTGACCTGCTCCATTTCTCACCTCAGGCTTCCAGCCCGCTCAGCATCAAGC 1320
DB 1261 ACTCAGATGTGGCTGACCTGCTCCATTTCTCACCTCAGGCTTCCAGCCCGCTCAGCATCAAGC 1320

QY 1321 CCAGCAAAAGCAGACAAACCCCATCTGCTCACCATCCAGCCGGACCTCAGCACCCACCA 1380
DB 1321 CCAGCAAAAGCAGACAAACCCCATCTGCTCACCATCCAGCCGGACCTCAGCACCCACCA 1380

QY 1381 CCTACAGGCGAGTCTCTGTCCCGGAGGATGGGCCAGCCCAAGATTCAGCTCAGCTCA 1440
DB 1381 CCTACAGGCGAGTCTCTGTCCCGGAGGATGGGCCAGCCCAAGATTCAGCTCAGCTCA 1440

QY 1441 ATGGGCACTGTCTCAGCCCGCTGGGTCGCGCCGACACACTGTGCACACAGCTCTCCCA 1500
DB 1441 ATGGGCACTGTCTCAGCCCGCTGGGTCGCGCCGACACACTGTGCACACAGCTCTCCCA 1500

QY 1501 CCTCTGAGCCGAGGAGTTCGTCTCCCGCTCTCCACCCAGAACTACTTCCGCTCCCTGC 1560
DB 1501 CCTCTGAGCCGAGGAGTTCGTCTCCCGCTCTCCACCCAGAACTACTTCCGCTCCCTGC 1560

QY 1561 CCGGAGGCGACAGCAACATGACCTATGGGACCTTCAACTTCTCGGGGCGGCTGATGA 1620
DB 1561 CCGGAGGCGACAGCAACATGACCTATGGGACCTTCAACTTCTCGGGGCGGCTGATGA 1620

QY 1621 TCCCTAATACAGGTATCAGCTCTCTATCCCGCCAGATGCCATACCCCGAGGGAAGATCT 1680
DB 1621 TCCCTAATACAGGTATCAGCTCTCTATCCCGCCAGATGCCATACCCCGAGGGAAGATCT 1680

QY 1681 ATGAGATCTACCTCAGCTGCACAAAGCCGGAAGAGTGTGAGTTCCTTGTGCTGTCT 1740
DB 1681 ATGAGATCTACCTCAGCTGCACAAAGCCGGAAGAGTGTGAGTTCCTTGTGCTGTCT 1740

QY 1741 AGACCTGTGTAGTCCATCGT 1800
DB 1741 AGACCTGTGTAGTCCATCGT 1800

QY 1801 TCATCTGT 1860
DB 1801 TCATCTGT 1860

QY 1861 AAAAGCAGT 1920
DB 1861 AAAAGCAGT 1920

QY 1921 CCTTCCACCTCTACTACTGCGAGCTGAGGCGGCTGCTGCTGCTCTTTCACCCGAGCAGC 1980
DB 1921 CCTTCCACCTCTACTACTGCGAGCTGAGGCGGCTGCTGCTGCTCTTTCACCCGAGCAGC 1980

Qy	1981	TGGGCGGCTTTGGCTGGTGGAGAGGCCCTCAGCGTGGCTGCCGCAAGCGCTCAAGC	2040
Db	1981	TGGGCGGCTTTGGCTGGTGGAGAGGCCCTCAGCGTGGCTGGCGCAAGCGCTCAAGC	2040
Qy	2041	TGCTTCTGTTTGGGCGGTGCGCTGCACCTCCCTCGAGTACAAATCCGGGTCTACTGCC	2100
Db	2041	TGCTTCTGTTTGGGCGGTGCGCTGCACCTCCCTCGAGTACAAATCCGGGTCTACTGCC	2100
Qy	2101	TGCATGACACCCACGATGSCACTCAAGGAGTGGTGCAGCTGGAGAAAGACGTGGGGGGAC	2160
Db	2101	TGCATGACACCCACGATGSCACTCAAGGAGTGGTGCAGCTGGAGAAAGACGTGGGGGGAC	2160
Qy	2161	AGCTGATCCAGGAGCCACGGGTCCCTGCACCTTCAAGGACAGTTACCAAACTCGCCCTAT	2220
Db	2161	AGCTGATCCAGGAGCCACGGGTCCCTGCACCTTCAAGGACAGTTACCAAACTCGCCCTAT	2220
Qy	2221	CCATCCAGGATGTGCCAGCTCCCTGTGGAGAGTAAGTCTCTTGTGAGTACACGAGGA	2280
Db	2221	CCATCCAGGATGTGCCAGCTCCCTGTGGAGAGTAAGTCTCTTGTGAGTACACGAGGA	2280
Qy	2281	TCCCTCTTTTATCATACTCGGAATGCGACGACGGTACTTGCATGCACCTTACCCCTGG	2340
Db	2281	TCCCTCTTTTATCATACTCGGAATGCGACGACGGTACTTGCATGCACCTTACCCCTGG	2340
Qy	2341	AGCGTGTCAGCCCCAGCACTAGTGACCTGGCTGCAAGCTGTGGGTGTCAGCGTGGAGG	2400
Db	2341	AGCGTGTCAGCCCCAGCACTAGTGACCTGGCTGCAAGCTGTGGGTGTCAGCGTGGAGG	2400
Qy	2401	GCAGCGGCGAGAGCTTCAGCATCAACTTCAACATCAACAAAGGACAAAGTTTGTCTGAGC	2460
Db	2401	GCAGCGGCGAGAGCTTCAGCATCAACTTCAACATCAACAAAGGACAAAGTTTGTCTGAGC	2460
Qy	2461	TGCTGGCTCTGGAGAGTGAACGGGGGTCCAGCCCTGTGGTGGCGCCAGTGCCTTCAAGA	2520
Db	2461	TGCTGGCTCTGGAGAGTGAACGGGGGTCCAGCCCTGTGGTGGCGCCAGTGCCTTCAAGA	2520
Qy	2521	TCGCCCTTCCTCATTTGGSCAGAAGATAATTTCCAGCCTGGACCCACCTGTAGCGGGGTG	2580
Db	2521	TCGCCCTTCCTCATTTGGSCAGAAGATAATTTCCAGCCTGGACCCACCTGTAGCGGGGTG	2580
Qy	2581	CCGACTGGCGGACTCTGGGCCCAGAACTCCACTGGACAGCCATCTCAGCTTCTTTTGCT	2640
Db	2581	CCGACTGGCGGACTCTGGGCCCAGAACTCCACTGGACAGCCATCTCAGCTTCTTTTGCT	2640
Qy	2641	CCAAGCCACGCCACACAGCCATCATCTCAACTGTGGAGGGCGGCACTTCCCAACG	2700
Db	2641	CCAAGCCACGCCACACAGCCATCATCTCAACTGTGGAGGGCGGCACTTCCCAACG	2700
Qy	2701	GCAACTTCAGCCAGCTGGCTGCAGCAGTGGCTGGACTGGGCCACGACAGCTGGCTCT	2760
Db	2701	GCAACTTCAGCCAGCTGGCTGCAGCAGTGGCTGGACTGGGCCACGACAGCTGGCTCT	2760
Qy	2761	TCACAGTGTGGAGGCTGAGTGTGAGGGCGGCCAGGCCCGACACTACACTCTCACCCAG	2820
Db	2761	TCACAGTGTGGAGGCTGAGTGTGAGGGCGGCCAGGCCCGACACTACACTCTCACCCAG	2820
Qy	2821	CTTTGGCACCCACCAAGGACAGGCAGAACCGGACAGGGGCCCTTCCCCACACCGGGAG	2880
Db	2821	CTTTGGCACCCACCAAGGACAGGCAGAACCGGACAGGGGCCCTTCCCCACACCGGGAG	2880
Qy	2881	A 2881	
Db	2881	A 2881	

RESULT, T 2

RESUL 2
ABK37922

ID ABK37922 standard; cDNA; 2752 BP.

XX

AC
ABK

cdDNA encoding Human protein NOV1.

Human; NOVX; ss; gene; cardiomyopathy; atherosclerosis; diabetes; cell signal processing disorder; metabolic disorder; obesity; infection; anorexia; cancer-associated cachexia; cancer; neurodegenerative disorder; Alzheimer's disease; Parkinson's disease; immune disorder; haematopoietic disorders; dyslipidaemia; pain; asthma; hypertension; osteoporosis; Crohn's disease; multiple sclerosis; angina pectoris; myocardial infarction; ulcer; allergy; benign prostatic hypertrophy; psychosis; neurological disorder; anxiety; schizophrenia; manic depression; dementia; dyskinesia; Huntington's disease; Gilles de la Tourette's syndrome; gene therapy.

Homo sapiens.

WO200210216-A2.

07-FEB-2002.

30-JUL-2001: 2001WO-US024225.

28-JUL-2000; 2000US-0221409P;
04-AUG-2000; 2000US-0222840P;
04-AUG-2000; 2000US-0223752P;
04-AUG-2000; 2000US-0223762P;
04-AUG-2000; 2000US-0223769P;
04-AUG-2000; 2000US-0223770P;
14-AUG-2000; 2000US-0221745P;
15-AUG-2000; 2000US-0225392P;
15-AUG-2000; 2000US-0225470P;
16-AUG-2000; 2000US-0225697P;
01-FEB-2001; 2000US-0236362P;
05-APR-2001; 2001US-0281645P;

(CURA-) CURAGEN CORP.

Padigaru M, Mezes P, Mishra V, Burgess C, Casman S, Grosse WM, Alsbrook JP, Jephley DM, Gerlach VL, Macdougall JR, Smithson G, WPI; 2002-180074/23.
P-PSDB: AAU85403.

New isolated cytoplasmic, nuclear, membrane bound, or secreted polypeptide, useful for treating cardiomyopathy, atherosclerosis, infarctions, cancer, neurodegenerative, metabolic, hematopoietic and immune disorders.

Claim 9; Page 9-10; 213pp: English.

The invention relates to an isolated cytoplasmic, nuclear, membrane bound, or secreted polypeptide (NOVX, x= 1-14) their variants or mature form. Also included are the nucleic acids encoding the NOVX proteins, a vector comprising the nucleic acid, a cell comprising the vector, an anti-NOVX antibody and modulators of NOVX. NOVX, the NOVX-associated antibody are useful for treating or preventing a NOVX-associated disorder, where the disorder is selected from cardiomyopathy, atherosclerosis, diabetes, a disorder related to cell signal processing and metabolic pathway modulation, metabolic disorders, obesity, infectious disease, anorexia, cancer-associated cachexia, cancer, neurodegenerative disorders, Alzheimer's disease, Parkinson's disease, immune disorders, haematopoietic disorders, and the various dyslipidaemias, metabolic disturbances associated with obesity, the metabolic syndrome X and wasting disorders associated with chronic diseases, bacterial, fungal, protozoal and viral infections, pain, bulimia, asthma, hypertension, urinary retention, osteoporosis, Crohn's disease, multiple sclerosis, Albright Hereditary Osteodystrophy, angina pectoris, myocardial infarction, ulcer, allergy, benign prostatic hypertrophy, and psychotic and neurological disorders, including anxiety, schizophrenia, manic depression, delirium, dementia, and dyskinesias, such as Huntington's disease and Gilles de la Tourette's syndrome. The nucleic acid is useful in gene therapy. The present sequence encodes a NOVX protein

[illegible]

QY	567	TCCCTGAGCAGGGGATCGTGTGCCCTGCGTGCATCCACCGAGGGGATCCCTCCACCGGAG	626	QY	1644	CTCATCCCCCAGATGCCATATACCCGAGGGAGATCTATGAGATCTACTCATCGTGCAC	1703
Db	481	TCCCTGAGCAGGGGATCGTGTGCCCTGCGTGCATCCACCGAGGGGATCCCTCCACCGGAG	540	Db	1558	CTCATCCCCCAGATGCCATATACCCGAGGGAGATCTATGAGATCTACTCATCGTGCAC	1617
QY	627	GTGGAGTGGCTCCGGAACGAGACCTGTGTGACCGCTCCCTGACCCCAATATATACATC	686	QY	1704	AAGCGGAGACGTGAGGTTGCCCTAGCTGTGTGAGCCCTCTGAGTCCCATCGTT	1763
Db	541	GTGGAGTGGCTCCGGAACGAGACCTGTGTGACCGCTCCCTGACCCCAATATATACATC	600	Db	1618	AAGCGGAGACGTGAGGTTGCCCTAGCTGTGTGAGCCCTCTGAGTCCCATCGTT	1677
QY	687	ACGCGGAGCAGAGCTGTGTGTGAGCAGGGCCCGCTTGTGTGACACGGCCCAACTACACC	746	QY	1764	AGCTGTGACCCCTGCGGTCTGCTCACCGGCCAGTCTATCTGGCTATGACCACTGT	1823
Db	601	ACGCGGAGCAGAGCTGTGTGTGAGCAGGGCCCGCTTGTGTGACACGGCCCAACTACACC	660	Db	1678	AGCTGTGACCCCTGCGGTCTGCTCACCGGCCAGTCTATCTGGCTATGACCACTGT	1737
QY	747	TGCTGTGGCAAGAAATGTGTGCAAGTGTGCGCAGCGCTCCGCTGTGTATGCTTAC	806	QY	1824	GCGGAGCCAGCCCTGACAGCTGTGAGCCTTGCCCTCAAAAAGACGTGTGTGAGGGCAGC	1883
Db	661	TGCTGTGGCAAGAAATGTGTGCAAGTGTGCGCAGCGCTCCGCTGTGTATGCTTAC	720	Db	1738	GCGGAGCCAGCCCTGACAGCTGTGAGCCTTGCCCTCAAAAAGACGTGTGTGAGGGCAGC	1797
QY	807	GTGAACGGTGGTGTGACGTGGACCGAGTGGTCCGCTGTGAGCGCCAGCTGTGGCGC	866	QY	1884	TGGGAGCAGGATGTCTGCACTGTGGCGAGAGGGCCCTCCCACTCTACTACTGCCAG	1943
Db	721	GTGAACGGTGGTGTGACGTGGACCGAGTGGTCCGCTGTGAGCGCCAGCTGTGGCGC	780	Db	1798	TGGG---AGGATGTCTGACCTGTGGCGAGGAGGGCCCTCCCACTCTACTACTGCCAG	1854
QY	867	GGCTGGCAGAAACGAGCGGAGCTGCACCAACCCGGCGCCTCTCAACGGGGCGCTTTC	926	QY	1944	CTGGAGGCCAGTCCCTGCTTACCTGTTTCAACCGAGCAGCTGGGCGCTTGGCCCTGGTGGGA	2003
Db	781	GGCTGGCAGAAACGAGCGGAGCTGCACCAACCCGGCGCCTCTCAACGGGGCGCTTTC	840	Db	1855	CTGGAGGCCAGTCCCTGCTTACCTGTTTCAACCGAGCAGCTGGGCGCTTGGCCCTGGTGGGA	1914
QY	927	TGTGAGGGCAGATGTTCATGACCGCAGCGTCTCCTCTCTGCTGTCTGTGACGGC	986	QY	2004	GAGGCGCTCAGCGTGGCTGCCCGCAAGCGCCTCAAGCTGCTTCTGTGTGCGCGGTGGCC	2063
Db	841	TGTGAGGGCAGATGTTCATGACCGCAGCGTCTCCTCTCTGCTGTCTGTGACGGC	897	Db	1915	GAGGCGCTCAGCGTGGCTGCCCGCAAGCGCCTCAAGCTGCTTCTGTGTGCGCGGTGGCC	1974
QY	987	ASCTGAGCCCTGGAGCAAGTGGTGGGCTGTGAGCTGTGAGCCCACTGGCGAGC	1046	QY	2064	TGCACCTCCTCGAGTAAACATTCGGGTCTACTGCTGCTGATGACACCACTGACATC	2123
Db	898	ASCTGAGCCCTGGAGCAAGTGGTGGGCTGTGAGCTGTGAGCCCACTGGCGAGC	957	Db	1975	TGCACCTCCTCGAGTAAACATTCGGGTCTACTGCTGCTGATGACACCACTGACATC	2034
QY	1047	CGTGTGTCTGTACCAAGCAACCCGCAACGAGGGGAGGTGCCAGGGGCACTGACCTG	1106	QY	2124	AAGGAGGTGCTCAGCTGGAGAAAGCAGCTGGGGGACAGCTGATCCAGAGGCCACGGGTG	2183
Db	958	CGTGTGTCTGTACCAAGCAACCCGCAACGAGGGGAGGTGCCAGGGGCACTGACCTG	1017	Db	2035	AAGGAGGTGCTGAGCTGGAGAAAGCAGCTGGGGGACAGCTGATCCAGAGGCCACGGGTG	2094
QY	1107	GACACCGCAACTGTACAGTACCTCTGTGTACAGTGTCTTGGCCCTGAGGACGTG	1166	QY	2184	CTGCACCTCAAGGACAGTTTACCAACCTGCGCCTATCCATCCAGATGTGCCAGCTCC	2243
Db	1018	GACACCGCAACTGTACAGTACCTCTGTGTACAGTGTCTTGGCCCTGAGGACGTG	1077	Db	2095	CTGCACCTCAAGGACAGTTTACCAACCTGCGCCTATCCATCCAGATGTGCCAGCTCC	2154
QY	1167	GCCTCTATGTGGGCTCATCGCGTGGCGCTGTGCTGTGCTGTGCTGTGCTGTGCTC	1226	QY	2244	CTGTGGAAGAGTAAGTCTCTGTGACCTACAGGAGATCCCTTTTATCAGATCTGGAAT	2303
Db	1078	GCCTCTATGTGGGCTCATCGCGTGGCGCTGTGCTGTGCTGTGCTGTGCTGTGCTC	1137	Db	2155	CTGTGGAAGAGTAAGTCTCTGTGACCTACAGGAGATCCCTTTTATCAGATCTGGAAT	2214
QY	1227	ATCTCTGTTTATTCGCGAAGAGAGGGGCTGGAATCAGATGTGGCTGACTGCTCAT	1286	QY	2304	GGCAGCAGCGGTACTTGCACCTTGCACCTGGAGGTGTGAGCCCGCAGCACTAGT	2363
Db	1138	ATCTCTGTTTATTCGCGAAGAGAGGGGCTGGAATCAGATGTGGCTGACTGCTCAT	1197	Db	2215	GGCAGCAGCGGTACTTGCACCTTGCACCTGGAGGTGTGAGCCCGCAGCACTAGT	2274
QY	1287	CTCAGCTCAGGCTTCCAGCCGTCAGCATCAAGCCGAGCAAGCAGACACCCCACTG	1346	QY	2364	GACCTGGCCTGCAAGCTGTGGGTGTGGCAGGTGGAGGGGAGCGGGCAGAGCTTCAGCATC	2423
Db	1198	CTCAGCTCAGGCTTCCAGCCGTCAGCATCAAGCCGAGCAAGCAGACACCCCACTG	1257	Db	2275	GACCTGGCCTGCAAGCTGTGGGTGTGGCAGGTGGAGGGGAGCGGGCAGAGCTTCAGCATC	2334
QY	1347	CTCAGCTCAGGCTTCCAGCCGTCAGCATCAAGCCGAGCAAGCAGACACCCCACTG	1403	QY	2424	AACTTCAACATCACCAAGGACACAGGTTTGTGAGCTGCTGAGTGTGAGAGCG	2483
Db	1258	CTCAGCTCAGGCTTCCAGCCGTCAGCATCAAGCCGAGCAAGCAGACACCCCACTG	1317	Db	2335	AACTTCAACATCACCAAGGACACAGGTTTGTGAGCTGCTGAGTGTGAGAGCG	2394
QY	1404	CGGAGGATGGGCCAGCCCAAGTTCAGCTCAACATGGGCACTGTGAGCCCGCTG	1463	QY	2484	GGGTCCTCAGCCCTGTGTGGGCCCTCAGTGCCTTCAAGATCCCTTCTCATTCGCGAGAAG	2543
Db	1318	CGGAGGATGGGCCAGCCCAAGTTCAGCTCAACATGGGCACTGTGAGCCCGCTG	1377	Db	2395	GGGTCCTCAGCCCTGTGTGGGCCCTCAGTGCCTTCAAGATCCCTTCTCATTCGCGAGAAG	2454
QY	1464	GGTGGCGCGCCACACCTGACACAGCTCTCCACCTCTCAGGCGGAGGAGTTGCTC	1523	QY	2544	ATAAATTCAGCTGACCCACCTGTAGCGGGGTGCGGACTGTGGCGGAGCTTCGGCCCGAG	2603
Db	1378	GGTGGCGCGCCACACCTGACACAGCTCTCCACCTCTCAGGCGGAGGAGTTGCTC	1437	Db	2455	ATAAATTCAGCTGACCCACCTGTAGCGGGGTGCGGACTGTGGCGGAGCTTCGGCCCGAG	2514
QY	1524	TCCCGCTCTCCACCCAGAACTACTTCCGCTCCCTGCTCCCGGAGCAGCAACATGACC	1583	QY	2604	AAACTCCAGCTGACAGCCATCTCAGCTCTTTTGGCTTCCAAAGCCAGCCACAGGCCATG	2663
Db	1438	TCCCGCTCTCCACCCAGAACTACTTCCGCTCCCTGCTCCCGGAGCAGCAACATGACC	1497	Db	2515	AAACTCCAGCTGACAGCCATCTCAGCTCTTTTGGCTTCCAAAGCCAGCCACAGGCCATG	2574
QY	1584	TATGGGACCTTCACTCTCGGGGCGCGGTGATGATCCCTTAATACAGGTATCAGCCTC	1643	QY	2664	ATCCTCAAACCTGTGGGAGCGCGGCACTTCCCAACGGCAACCTCAGCCAGCTGGCTGCA	2723
Db	1498	TATGGGACCTTCACTCTCGGGGCGCGGTGATGATCCCTTAATACAGGTATCAGCCTC	1557	Db	2575	ATCCTCAAACCTGTGGGAGCGCGGCACTTCCCAACGGCAACCTCAGCCAGCTGGCTGCA	2634
				QY	2724	GCAGTGGCTGGAATGGGCGAGCCAGAGCTGGGCTCTTACAGTGTGCGAGGCTGAGTGC	2783

```

Db 2635 GCAGTGGCTGGACTGGCGCAGCAGCTGGCTCTTCACATGTCGGAGGCTGAGTGC 2694
QY 2784 TGA 2786
Db 2695 TGA 2697

RESULT 4
ABK15169
ID ABK15169 standard; DNA; 3580 BP.
XX
AC ABK15169;
XX
DT 23-APR-2002 (first entry)
XX
DE Human REPTR 1 cDNA sequence.
XX
KW REPTR; human; antiinflammatory; cytostatic; immunosuppressive; antiviral;
KW anti-HIV; antiarthritic; anticonvulsant; nootropic; neuroprotective;
KW antiallergic; antibody; immunogen; endometriosis;
KW gastrointestinal disorder; gastritis; oesophageal carcinoma;
KW Crohn's disease; irritable bowel syndrome; ulcerative colitis;
KW endocrine disorder; hypothyroidism; Kallman's disease;
KW autoimmune disease; inflammatory disease; infertility; receptor;
KW acquired immune deficiency syndrome; AIDS; rheumatoid arthritis; allergy;
KW osteoarthritis; diabetes mellitus; multiple sclerosis;
KW systemic lupus erythematosus; cell proliferative disorder; cancer;
KW developmental disorder; Duchenne muscular dystrophy; gene;
KW Becker muscular dystrophy; neurological disorder; epilepsy;
KW Alzheimer's disease; Huntington's disease; reproductive disorder; ss.
XX
OS Homo sapiens.
XX
PH Key Location/Qualifiers
FT CDS 4..2532
FT /*tag= a
FT /product= "REPTR1 protein"
XX
PN WO200198354-A2.
XX
PD 27-DEC-2001.
XX
PF 21-JUN-2001; 2001WO-US019942.
XX
PR 21-JUN-2000; 2000US-0214027P.
PR 25-AUG-2000; 2000US-0228045P.
PR 12-DEC-2000; 2000US-0255104P.
XX
EA (INCY-) INCYTE GENOMICS INC.
XX
PI Griffin JA, Kallick DA, Tribouley CM, Yue H, Nguyen DB, Tang YT;
PI Lal P, Policky JL, Aizimai Y, Lu DAM, Graul R, Yao MG, Burford N;
PI Hafalia AJA, Baughn MR, Bandman O, Patterson C, Yang J, Xu Y;
PI Gandhi AR, Warren BA, Ding L, Sanjanwala MS, Duggan BM, Lu Y;
XX
DR WPI; 2002-090432/12.
DR P-PSDB; AAU17818.
XX
PT Twelve human receptors (referred to as REPTR-1 to REPTR-12), useful in
PT the diagnosis, treatment and prevention of gastrointestinal (e.g.
PT gastritis), autoimmune/inflammatory (e.g. osteoarthritis) and cell
PT proliferative (e.g. cancer) disorders.
XX
PS Claim 57; Page 142-143; 157pp; English.
XX
CC This invention relates to twelve human receptors cDNA sequences referred
CC to as REPTR-1 to REPTR-12, and the proteins encoded thereby. The
CC proteins of the invention may have antiinflammatory, cytostatic,
CC immunosuppressive, antiviral, anti-HIV, antiarthritic, muscular active
CC general, anticonvulsant, nootropic, neuroprotective, antiallergic
CC activities. The sequences of the invention may be used to produce REPTR
CC agonists or antagonists, and the protein sequences may be used to raise

```

```

CC anti-REPTR antibodies. These molecules and the REPTR polynucleotides and
CC polypeptides of the invention are useful in the diagnosis, treatment and
CC prevention of gastrointestinal (e.g. gastritis, oesophageal carcinoma,
CC Crohn's disease, irritable bowel syndrome, ulcerative colitis), endocrine
CC (e.g. hypothyroidism, Kallman's disease), autoimmune/inflammatory
CC (e.g. acquired immune deficiency syndrome (AIDS), rheumatoid arthritis,
CC allergies, osteoarthritis, diabetes mellitus, multiple sclerosis,
CC systemic lupus erythematosus), cell proliferative (e.g. cancer),
CC developmental (e.g. Duchenne and Becker muscular dystrophy), neurological
CC (e.g. epilepsy, Alzheimer's disease, Huntington's disease) and
CC reproductive (e.g. infertility, endometriosis) disorders. Numerous other
CC examples of each disorder are given in the specification. The present
CC sequence represents the human REPTR1 cDNA sequence of the invention
XX
SQ Sequence 3580 BP; 670 A; 1233 C; 1025 G; 652 T; 0 U; 0 Other;

```

```

Query Match 83.4%; Score 2402.6; DB 6; Length 3580;
Best Local Similarity 93.4%; Pred. No. 0;
Matches 2615; Conservative 0; Mismatches 9; Indels 177; Gaps 3;
QY 84 GCATGGCGCTCCGGCCCGGCTGTGGCCAGCGCTCTGGGCATAGTCTCCGCGCTGG 143
Db 1 GCATGGCGCTCCGGCCCGGCTGTGGCCAGCGCTCTGGGCATAGTCTCCGCGCTGG 60
QY 144 CTCGGCGCTCGGTGGCCAGCAGAGTGCCACCGTGGCCAAACCCAGTGCTGGTGCCAA 203
Db 61 CTCGGCGCTCGGTGGCCAGCAGAGTGCCACCGTGGCCAAACCCAGTGCTGGTGCCAA 120
QY 204 CCGGACCTGCTTCCCACTTCTCTGTGGAGCCCGAGGATGTATCATGTCGAAGAACAG 263
Db 121 CCGGACCTGCTTCCCACTTCTCTGTGGAGCCCGAGGATGTATCATGTCGAAGAACAG 180
QY 264 CCAGTGCTGTGTGTGCAAGCCCGTGGCCAGCAGATCTTCTTCAAGTGCAACGGG 323
Db 181 CCAGTGCTGTGTGTGCAAGCCCGTGGCCAGCAGATCTTCTTCAAGTGCAACGGG 240
QY 324 GAGTGGTGCGCCAGGTGGACCACTGATGAGCGCAGCAGCGGAGCAGTGGTGAG 383
Db 241 GAGTGGTGCGCCAGGTGGACCACTGATGAGCGCAGCAGCGGAGCAGTGGTGAG 300
QY 384 CCGACCATGGAGTCCGCAATTAATGCTCTCAAGGACAGGTTCGAGAGGTGTTCGGGCTG 443
Db 301 CCCACCATGGAGTCCGCAATTAATGCTCTCAAGGACAGGTTCGAGAGGTGTTCGGGCTG 360
QY 444 GAGGAATACTGTCGCGAGTGGTGGCATGAGCTCTCGGCACCCACCAAGAGTCAGAG 503
Db 361 GAGGAATACTGTCGCGAGTGGTGGCATGAGCTCTCGGCACCCACCAAGAGTCAGAG 420
QY 504 GCCTACATCCGCATAGCCAGATTTCGCAAGAACTTCGAGCAGGAGCCGCTGGCCAAAGAG 563
Db 421 GCCTACATCCGCATAGCCAGATTTCGCAAGAACTTCGAGCAGGAGCCGCTGGCCAAAGAG 480
QY 564 GTGTCCCTGGAGCAGGCGCATCGTCTGCGCTCCCGTCCACCGAGGAGCATCCCTCCAGCC 623
Db 481 GTGTCCCTGGAGCAGGCGCATCGTCTGCGCTCCCGTCCACCGAGGAGCATCCCTCCAGCC 540
QY 624 GAGTGGAGTGGTCCGGAACGAGGACCTGTGTGACCCCGTCCCTGGACCCCAATGTATAC 683
Db 541 GAGTGGAGTGGTCCGGAACGAGGACCTGTGTGACCCCGTCCCTGGACCCCAATGTATAC 600
QY 684 ATCAGCGGAGCAGCAGCTGGTGGTGGAGCAGGCGCGCTTGTGTGACAGGCGCAATAC 743
Db 601 ATCAGCGGAGCAGCAGCTGGTGGTGGAGCAGGCGCGCTTGTGTGACAGGCGCAATAC 660
QY 744 ACCTGCGTGGCCAAAGAACTTCGCGCAGCTCGCGCAGCGCTCCGCTGCTGTATGTC 803
Db 661 ACCTGCGTGGCCAAAGAACTTCGCGCAGCTCGCGCAGCGCTCCGCTGCTGTATGTC 720
QY 804 TAGCTGAACGCTGGTGGTGGTGGAGCAGGAGTGGTCCGCTTCGAGCGCGCAGCTGTGGG 863
Db 721 TAC-----
QY 864 CGCGGCTGGCAGAAACGGAGCCGGAGTGTCACCAACCCCGCGCTCTCAACCGGCGGCGCT 923

```


Human; cytokine; cell proliferation; cell differentiation; gene therapy; vaccine; peptide therapy; stem cell growth factor; haematopoiesis; tissue growth factor; immunomodulatory; cancer; leukaemia; nervous system disorder; arthritis; inflammation; ss.

Homo sapiens.

WO200157190-A2.

09-AUG-2001.

05-FEB-2001; 2001WO-US004098.

03-FEB-2000; 2000US-00496914.

27-APR-2000; 2000US-00560875.
20-JUN-2000; 2000US-00598075.

19-JUL-2000; 2000US-00620325.
01-SEP-2000; 2000US-00654936.

15-SEP-2000; 2000US-00663561.
20-OCT-2000; 2000US-00693325

30-NOV-2000; 2000US-00728422.

(HYSE-) HYSEQ INC.

Tang YT, Liu C, Drmanac RT, Ma Y, Zhao QX, Wang D, Wang

Xue AJ, Yang Y, Wejhrman T,

WPI; 2001-476283/5

100

in diagnosis and gen

Claim 1; Page 2691-2

The invention relates

cytokine, cell proliferation

production of other cytokines in other cell populations. The polynucleotides and polypeptides are useful in gene therapy, vaccines or peptide therapy. The polypeptides have various cytokine-like activities.

e.g. stem cell growth factor activity, ha
tissue growth factor activity

activin/inhibin activity and may be useful in the diagnosis and/or treatment of cancer, leukaemia, nervous system disorders, arthritis and inflammation. Note: Records for SEQ ID NO 2110 (AAK53581), 2111 (AAK52582) and 3666 (AAM80020) are omitted as the relevant pages from the sequence listing were missing at the time of publication

Sequence 2907 BP; 552 A; 966 C; 881 G; 508 T; 0 U; 0 Other;

Query Match	79.7%;	Score 2296.2;	DB 4;	Length 2907;
Best Local Similarity	92.0%;	Pred. No. 0;		
Matches 2565;	Conservative	0;	Mismatches 38;	Indels 186;
Gaps 6				

133 TCGCCGCTTGGCTCCGCGGCTCGGGTGCCAGCAGTGCCACCGTGCCAAACCCAGTGC 192

170 TGGCGGCAGACAGCACGGATGGAGCGCCAGCAGAGTGCCACCGTGCGCCAACCCAGTGC 229

193 CTGGTGCCAAACCCGGACCTGCTTCCCCACTTCCTGGTGAGCCCGAGGATGTGTACATCG 252

230 CTGGTGCCAACCCGGACCTGCTTCCCCACTTCCCTGGTGAGCCCCGAGGATGTGTACATCG 288

253 TCAAGAACAGCCAGTGCTTGTTGTGTGCAAGGCCGTGCCGCCACGCACGATCTTCTTCA 313

290 TCAAGAAC AAGCCAGT GCTGCTGT GTGGCA AGGCCGT GCCCGCC CACCATC CTTCCTTA

313 AGTGCAACGGGAGTGGGTGGCGCAGCTCA CCACTCATTCATCCGCACGCCG

[illegible][illegible]

Db	410	GCAGTGGGCTGCCCA	CTATGGAGGTCCGCA	TTAAATGTCTCAAGGCAGCAGGTCTGGAAGG	469
Qy	433	TGTTTCGGGCTGGAGGA	TACTTGTTGGCCAGTGGCATGGAGCTCTCTCGGGCACCCACA	492	
Db	470	TGTTTCGGGCTGGAGGA	TACTTGTTGGCCAGTGGCATGGAGCTCTCTCGGGCACCCACA	529	
Qy	493	AGAGTCAGAAAGGCTTAC	TCCTCCGATAGCCAGATTGGCAAGAACTTCGAGCAGGAGCCGC	552	
Db	530	AGAGTCAGAAAGGCTTAC	TCCTCCGATAGCCAGATTGGCAAGAACTTCGAGCAGGAGCCGC	589	
Qy	553	TGSCCAAGAGAGTGTCCT	TGGAGCAGGGCATCGTCTGCCCTGCTGCCGTCTCACCGGAGGGCA	612	
Db	590	TGSCCAAGAGAGTGTCCT	TGGAGCAGGGCATCGTCTGCCCTGCTGCCGTCTCACCGGAGGGCA	649	
Qy	613	TCCCTCCAGCCAGAGTGGAGT	GGCTCCGGAAACGAGGACCTGTGTGACCCGCTCCCTGGACC	672	
Db	650	TCCCTCCAGCCAGAGTGGAGT	GGCTCCGGAAACGAGGACCTGTGTGACCCGCTCCCTGGACC	709	
Qy	673	CCAAATGTAATACACGCGGGAGCA	CAGCCTGGTGGCAGACAGGCCCGCCTTGCTGACA	732	
Db	710	CCAAATGTAATACACGCGGGAGCA	CAGCCTGGTGGCAGACAGGCCCGCCTTGCTGACA	769	
Qy	733	CGSCCAACTACACCTCGCTGGCCAGACAT	CTGTGGCAGCTGCGCGCAGCGCCTCCGCTG	792	
Db	770	CGSCCAACTACACCTCGCTGGCCAGACAT	CTGTGGCAGCTGCGCGCAGCGCCTCCGCTG	829	
Qy	793	CTGTCAATCGTCTTAC	-----	806	
Db	830	CTGTCAATCGTCTTAC	CGGTGGGCCCCGGGACTCCCTGGTCAAGGAGAGGACATCGCGGTGC	889	
Qy	807	-----	-----GTCAA	CGGTGGGCTGCTCGA	825
Db	890	CCCTGGGCGAGTGA	CATGTGCGTCTCTGTCTGTCGCGCCAGTGAACGTTGGGTGGTCTCGA	949	
Qy	826	CGTGCACCGAGTGTCCGTCTG	CACGCGCAGCTGTGGCGCGCGCTGGGAGAAACGAGACC	885	
Db	950	CGTGCACCGAGTGTCCGTCTG	CACGCGCAGCTGTGGCGCGCGCTGGGAGAAACGAGACC	1009	
Qy	886	GGAGCTGCACCAACCCCGCGCCTCT	CAACGGGGCGCTTTCTGTGAGGGCAGAAATGTC	945	
Db	1010	GGAGCTGCACCAACCCCGCGCCTCT	CAACGGGGCGCTTTCTGTGAGGGCAGAAATGTC	1069	
Qy	946	ATGACCGCACCGTCTCTCTCTG	TTGTTCTGTGTGACCGCAGCTGGAGCCCGTGGAGCA	1005	
Db	1070	AGAA-----AACAGCCTGCGCCAC	CCCTGTGCCAGTGGACGCGCAGCTGGAGCCCGTGGAGCA	1126	
Qy	1006	AGTGGTCGGCTGTGGGCTG	GAATGACACCACTGGCGGAGCGGTGAGTGTCTGACCCAG	1065	
Db	1127	AGTGGTCGGCTGTGGGCTG	GAATGACACCACTGGCGGAGCGGTGAGTGTCTGACCCAG	1186	
Qy	1066	CACCCCGCAACGAGGGGAGAGT	GGCGCAGGGCACTGACCTGGACACCCGCAACTGTACCA	1125	
Db	1187	CACCCCGCAACGAGGGGAGAGT	GGCGCAGGGCACTGACCTGGACACCCGCAACTGTACCA	1246	
Qy	1126	GTGACCTCTGTGTACACA	-----	1143	
Db	1247	GTGACCTCTGTGTACACA	CTCTACACCCCTGCCCCACCAAGGCCATGCTGTCTCCCG	1306	
Qy	1144	---GTGCTTCTGGCCCTGAGGA	GTGGCCCTCTATGTGGGCTCTATCGCGGTGGCCGTCT	1200	
Db	1307	CAGCTGTTCTTGGCCCTGAGGAG	GTGGCCCTCTATGTGGGCTCTATCGCGGTGGCCGTCT	1366	
Qy	1201	GCTTGTCTCTGCTGCTTGTCT	CTCATCTCTGTTTATTGCGGAGAGAGGAGGGCTGG	1260	
Db	1367	GCTTGTCTCTGCTGCTTGTCT	CTCATCTCTGTTTATTGCGGAGAGAGGAGGGCTGG	1426	
Qy	1261	ACTCAGATGTGGCTGA	CTCTGCTTCTCACTTCAGGCTTCAGACCGCTCAGCATCAAGC	1320	
Db	1427	ACTCAGATGTGGCTGA	CTCTGCTTCTCACTTCAGGCTTCAGACCGCTCAGCATCAAGC	1486	
Qy	1321	CCAGCAAGCAGACACCCCCATCT	GTCTCAACCATCAGCCGACCTCAG---CACCAACA	1377	

Db 1487 CCAGCAAGAGAGACAAACCCCATCTGCTCACCATCCAGCCGAGACCTCAGACCCACCACA 1546
 Qy 1378 CCACCTACAGGGAGTCTCTGTCCCGGCGAGGATGGGCCCAGCCAGCCCAAGTTCCAGTCA 1437
 Db 1547 CCACCTACAGGGAGTCTGTCTCCCGGCGAGGATGGGCCCAGCCAGTTCAGTCA 1606
 Qy 1438 CCAATGGGACCTGTCTAGCCCTCTGGTGGGCGGCGCACACACTGACACACAGCTCTC 1497
 Db 1607 CCAATGGGACCTGTCTAGCCCTCTGGTGGGCGGCGCACACACTGACACACAGCTCTC 1666
 Qy 1498 CCACCTCTGAGCGCGAGGAGTTCTCTCCCGCTCTCCACCCAGAACTACTTCCGCTCC 1557
 Db 1667 CCACCTCTGAGCGCGAGGAGTTCTCTCCCGCTCTCCACCCAGAACTACTTCCGCTCC 1726
 Qy 1558 TGCCCCGAGGACCCAGCAACATGACCTATGGACCTTCAACTTCTTGGGGGCGCGCTGA 1617
 Db 1727 TGCCCCGAGGACCCAGCAACATGACCTATGGACCTTCAACTTCTTGGGGGCGCGCTGA 1786
 Qy 1618 TGATCCCTAATACAGGTATCAGCTCTCCTATCCCCAGATGCCATACCCCGAGGGAAGA 1677
 Db 1787 TGATCCCTAATACAGGAATCAGCTCTCTATCCCCAGATGCCATACCCCGAGGGAAGA 1846
 Qy 1678 TCTATGAGATCTACCTCAGCTGCAACAGCGGAGACCTGAGGTTGCCCTAGCTGGCT 1737
 Db 1847 TCTATGAGATCTACCTCAGCTGCAACAGCGGAGACCTG----- 1887
 Qy 1738 GTCAGACCTGTGAGTCCCATCTGTTAGTGTGGACCCCTCTGGCGTCTCTACCCGCG 1797
 Db 1888 -----AGCTGTGGACCCCTCTGGCGTCTCTACCCGCG 1921
 Qy 1798 CAGTCATCTGTGATGAGGACCTGTGGGAGCCAGCCCTGACAGCTGAGGCTCGGC 1857
 Db 1922 CAGTCATCTGTGATGAGGACCTGTGGGAGCCAGCCCTGACAGCTGAGGCTCGGC 1981
 Qy 1858 TCAAAAGCAGTCTGTGAGGCGAGCTGGGAGGAGATGTGCTGCACTCTGGGCGAGGAG 1917
 Db 1982 TCAAAAGCAGTCTGTGAGGCGAGCTGGG--AGATGTGCTGCACTCTGGGCGAGGAG 2038
 Qy 1918 CGCCCTCCCACTCTACTGTCAGCTGAGGCGAGTGTGCTGCTAGCTTTACCGAGC 1977
 Db 2039 CGCCCTCCCACTCTACTGTCAGCTGAGGCGAGTGTGCTGCTAGCTTTACCGAGC 2098
 Qy 1978 AGTGGGCGCTTGTGCTGGGAGGAGCCCTCAGCGTGGCTGCGCCCAAGCGCTCA 2037
 Db 2099 AGTGGGCGCTTGTGCTGGGAGGAGCCCTCAGCGTGGCTGCGCCCAAGCGCTCA 2158
 Qy 2038 AGCTGCTTCTGTGCTGGGAGGAGGAGCTGCACTCTCGAGTACCAATCCGGTCTACT 2097
 Db 2159 AGCTGCTTCTGTGCTGGGAGGAGGAGCTGCACTCTCGAGTACCAATCCGGTCTACT 2218
 Qy 2098 GCTGCAATGACACCCAGTGCATCAAGAGGTGTGAGTGGAGAGCAGCTGGGG 2157
 Db 2219 GCTGCAATGACACCCAGTGCATCAAGAGGTGTGAGTGGAGAGCAGCTGGGG 2278
 Qy 2158 GACAGCTGATCCAGGAGCCAGGCTGCTGACCTCAAGGAGTACCAACACTGCGCC 2217
 Db 2279 GACAGCTGATCCAGGAGCCAGGCTGCTGACCTCAAGGAGTACCAACACTGCGCC 2338
 Qy 2218 TATCCATCCAGTGTGCCAGCTCTCTGTGGAGAGTAAAGCTCTTGTGAGTACAGG 2277
 Db 2339 TATCCATCCAGTGTGCCAGCTCTCTGTGGAGAGTAAAGCTCTTGTGAGTACAGG 2398
 Qy 2278 AGATCCCTTTATCAATCTGGAATGGACGCGGAGCTGCTGCACTGCACTTACCC 2337
 Db 2399 AGATCCCTTTATCAATCTGGAATGGACGCGGAGCTGCTGCACTGCACTTACCC 2458
 Qy 2338 TGAGGCTGTGAGCCAGCAGTGTGAGCTGCGCTGCAAGCTGTGGTGTGGCAGGTGG 2397
 Db 2459 TGGAGGCTGTGAGCCAGCAGTGTGAGCTGCGCTGCAAGCTGTGGTGTGGCAGGTGG 2518
 Qy 2398 AGGCGAGCGGAGAGCTTACGATCAACTTCAATCAATCAACAGGACACAAAGGTTGCTG 2457
 Db 2519 AGGCGAGCGGAGAGCTTACGATCAACTTCAATCAATCAACAGGACACAAAGGTTGCTG 2578

Qy 2458 AGCTGCTGCTCTGAGAGTGAAGCGGGGTCCAGCCCTGTGGGCCCCCAAGTGCCTTCA 2517
 Db 2579 AGCTGCTGCTCTGAGAGTGAAGCGGGGTCCAGCCCTGTGGGCCCCCAAGTGCCTTCA 2638
 Qy 2518 AGATCCCTTCTCTATTCGGCAGAAATAATTTCCAGCTGAGACCCACCTGTAGGCGG 2577
 Db 2639 AGATCCCTTCTCTATTCGGCAGAAATAATTTCCAGCTGAGACCCACCTGTAGGCGG 2698
 Qy 2578 GTGCCGACTGGGAGCTGTGGCCGAGAACTCCACTGTGAGAGCCATCTCAGCTTCTTTG 2637
 Db 2699 GTGCCGACTGGGAGCTGTGGCCGAGAACTCCACTGTGAGAGCCATCTCAGCTTCTTTG 2758
 Qy 2638 CCTCAAGCCAGCCACAGCCATGATCCTCAACTGTGGGAGGCGGCGCACTTCCCA 2697
 Db 2759 CCTCAAGCCAGCCACAGCCATGATCCTCAACTGTGGGAGGCGGCGCACTTCCCA 2818
 Qy 2698 ACGGCAACTCAGCAGCTGTGGCTGAGAGTGTGAGCTGAGCCAGCCAGAGCTGGCC 2757
 Db 2819 ACGGCAACTCAGCAGCTGTGGCTGAGAGTGTGAGCTGAGCCAGAGCCAGAGCTGGCC 2878
 Qy 2758 TCTTACAGTGTGGAGGCTGAGTCTGA 2786
 Db 2879 TCTTACAGTGTGGAGGCTGAGTCTGA 2907

RESULT 6

AAV52940

ID AAV52940 standard; cDNA; 3014 BP.

XX

AC AAV52940;

AC 25-MAR-2003 (revised)

DT 21-DEC-1998 (first entry)

XX

Rat UNC-5 homologue unc5h-1 cDNA.

DE

XX UNC-5; UNC5H-1; rat; netrin receptor; cell migration; axon guidance;

KW diagnosis; therapy; ds.

XX

Rattus sp.

XX

Key Location/Qualifiers

FT 1..2697

CDS /*tag= a

XX

MO9837085-A1.

XX

27-AUG-1998.

XX

19-FEB-1998; 98WO-US003143.

XX

19-FEB-1997; 97US-00808982.

XX

(REGC) UNIV CALIFORNIA.

XX

Tessier-Lavigne M, Leonardo ED, Hinck L, Masu M, Keinomasu K;

XX

WPI; 1998-495364/42.

XX

P-PSDB; AAW78898.

XX

Netrin-binding, vertebrate proteins - useful for diagnosis, therapy and

XX the biopharmaceutical industry.

XX

Claim 7; Page 15-17; 32pp; English.

XX

This cDNA, termed unc5h-1, comprises a rat homologue of Caenorhabditis
 elegans unc-5. Rat unc5h-1 and unc5h-2 (see AAV52942) cDNAs were isolated
 from an E18 brain cDNA library. The predicted proteins (see AAW78998 and
 AAW78900) show similarity with UNC-5. They are predicted to be involved
 in cell migration and axon guidance, and are characterised as receptor
 proteins for netrins. Gene expression is observed in regions where
 differentiating neurons are undergoing axogenesis. Human unc5h-1 (see

AAV52941) and unc5h-2 (see AAV52943) cDNAs are also claimed. Vertebrate UNC-5 proteins may be produced recombinantly from transfected host cells by utilising these vertebrate UNC-5 nucleic acids. The invention also provides unc-5 hybridisation probes and primers, vertebrate UNC-5-specific binding agents such as specific antibodies, and methods of making and using the subject compositions in diagnosis (e.g. genetic hybridisation screens for vertebrate unc-5 transcripts), therapy (e.g. gene therapy to modulate vertebrate unc-5 gene expression) and in the biopharmaceutical industry (e.g. as immunogens, reagents for modulating cell guidance, reagents for screening chemical libraries for lead pharmacological agents, etc.). (Updated on 25-MAR-2003 to correct PI field.)

SQ Sequence 3014 BP; 596 A; 977 C; 849 G; 592 T; 0 U; 0 Other;

Query Match	Score	DB 2	Length
77.6%	2235.8	DB 2	3014

Best Local Similarity 88.5%; Pred. No. 0;
Matches 2474; Conservative 0; Mismatches 312; Indels 11.

Matches 2474; Conservative 0; Mismatches 312; Indels 11; Gaps 4;

QY
87 ATGGCGGTCCGGCCCGGCGCTGTGGCCAGCGCTCCTGGGCATAGTCTCGCCGCTTGGCTC 146
+ + + + +

db
1 ATGCGCGTCCGGCCCGGCGCTGTGGCCAGTGTGTCCTGGGCA TAGTCTCTGCGCGCGCTCGGTT 60

[illegible][illegible]

61 CGGGTCCAGAGAGTGCCACGGTGGCCAATCCAGTGCCCGGTGCCAACCC 120

207 GACCTGCTTCCCCACTTCTGGTGGAGCCCGAGGATGTGTACATGCTCAAGACAAGCCA 266

Db 121 GACCTGCTGCCCCACTTCCTGGTAGAGCCTGAGGACGTTACATTGTCAAGAACAGCCG 180

Qy 267 GTGCTGCTTGTCGAAGCCGCTGCCCGCCACGCAGATCTTCTTCAAGTGCACCGGGAG 326

Db
181 GTGTTGTGGTGTGCAAGCTGTGCCTGCCACCCAGATCTTCTTCAAGTGCAATGGGAA 240

QY 327 TGGGTGCGCCAGGTGGACCAAGTGATCGTGGCGAGCAAGAGCGGAGCGAGTGGTACACCC 396

[illegible][illegible]

387 ACCAAGGAGGATCCGCAATAAAGTCTCAAGSCAGCAGGTTCGAGAAGGTGTTCCGGGCTGGAG 446

301 ACCATGGAGGTCGGTATCAACGTATCGAGGCAGCAGGTAGAGAAAGTGTTTGGGTGGAG 360

447 GAATACTGGTGCACGTGGTGGCATGGAGCTCTCGGGCACCAAGAGTCAAGGCC 506

Db
361 GAATACTGGTGGCACTGTGTGGCATGGAGCTCTCGGGTACCAACCAAAGTCGAAGGCC 420

QY
507 TACATCCGCATAGCCAGATTGCGCAAGAACTTCGAGCAGGAGCCGCTGGCCACGAGGTG 566

Db
421 TACATCCGGATTGCCTATTTCGGCAGAACTTTGAGCAGAGAGCCACTGGCCACGTAAGTC 480

567 TCCCTGGAGCAGGGCATCGTGTGCGCTGCGCTCCACCGCAACCCCATCCCGTCCACCGCG

[illegible]

501 TCACGGAGCAAGGCATGGTACCTGTGCGCCCCCAGAAGGAATCCCCCAGCTGAG 540

627 GTGGAGTGGTCCGGAACGAGGACCTGGTGGACCCGTCCTGGACCCCAATGTATACATC 686

Db
541 GTGGAGTGGCTCGAATGAGGACTGTGGACCCCTCCCTCGATCCCAATGTGTACATC 600

687 **ACGGGGAGCACAGCCTGGTGGTGGACAGGCCCGCCTTGCTGACACGGCACTACACC** 746

601 ACGGGGAGCACAGCCTAGTCGTGCGTCAGGCCCGCCTGGCCGACACGGGCCAATACACC 660

QY 747 TGGTGGCCAAAGAACATCGTGGGACGTGCGCGCAGGGGCTCCGCTGCTGCAATCTCTAC 806

[illegible]

097 CTCTTGGGAGGATGCGTCGACGGGTGCATTGTTAT 720

807 G G A C G G I G G G T G G T C G A C G T G G A C C G A G T G G T C C G T C T G C A G C G C C A G C T G T G G G C G C 866

Dbb 721 GTGAACGGTGGGTGGTGCAGCTGGACTGAGTGGTCCGTCTGCAGGCCAGCTGTGGGCGT 780

QY 867 GGCTGGCAGAAACGGAGCCGGAGCTGCACCAACCCGGCGCCTCTCAACGGGGCGCTTC 926

Db	781	GGCTGGCAGAAACGGAGCGCGAGCTGCACCAACCCGCGACCTCTCTCAAAGGGGGCGCCTTC	841
QY	927	TGTGAGGGGCGAAATGTCATCAGCCGCGCGCTCTCTCTGCTTGTCTCTGTGTGACAGCGC	986
Db	841	TGTGAGGGGCGAAGTTCAGNAACAGCGCTGGCCACTCTG---TGCCCAAGTGATGGG	897
QY	987	AGCTGGAGCCCGTGGAGCAAGTGTGCGCCTGTGGGCTGGACTGCACCCACTGGCGGAGC	1046
Db	898	AGCTGGAGTTGTTGAGTAAGTGTGAGCTGTGGCTTGACTGCACCCACTGGCGGAGC	957
QY	1047	CSTGTAGTGTCTGACCCAGCACCCCGGAAAGCGAGGGAGGAGTGCCAGGGCACTGACCTG	1106
Db	958	CGCGAGTGTCTGACCCAGCACCCCGAATGAGGTGAGGAGTGTGGGGTGTCTGACCTG	1017
QY	1107	GACACCCGCAACTGTACCAAGTACCTCTGTGTACACAGTGTCTTGGCCCTTGAGGAGTG	1168
Db	1018	GACACCCGCAACTGTACAGTGAACCTCTGCTGTGCACACCGCTTCTTGCCTCCGAGAGCGTG	1077
QY	1167	GGCCTCTATGTGGGCTCATCGCGGTGGCCGTCTGCCTGTGCTGTCTGTCTGTCTCTC	1226
Db	1078	GCTCTCTACATCGGCTTGTGCTGTGGCTGTGTGCTCTTCTTGTGTGTGTCTGCGCCTT	1137
QY	1227	ATCCTCGTTTATTCGGGAAAGAGGGGCTGTGACTCAGATGTGTGGCTGTACTGTCTCAAT	1286
Db	1138	GGACTCATTTACTGTGCGCAAGAGGAAGGGCTGACTCCGATGTGGCGCACTGTCTCCATC	1197
QY	1287	CTCAGCTCAGGCTTCCAGCCCGTCTCAGATCAAGCCAGCAAGAGAGAGCAACCCCCCATCTG	1346
Db	1198	CTCAGCTCGGCTTCCAGCTGTGAGCATCAAGCCCGAGAAAGCAGACAAACCCCACTG	1257
QY	1347	CTCAGCATTCAGCGGAGCTTCAGCAACACC---ACCACCTACCAAGGCGAGTCTCTGTCCC	1403
Db	1258	CTCACCATTCAGCGCAGACCTCAGCACCAACCACTACCACCTACCAAGGCGAGTCTATGTTG	1317
QY	1404	CGGCGAGATGGCGCCAGCGCCCAAGTTCAGCTCACAATGGGCACTGTCTAGCCCCCTG	1463
Db	1318	AGGCGAGATGACCCAGCGCCCAAGTTCAGCTCTCTAATGGTCACTGTCTCAGCCCACTG	1377
QY	1464	GGTGGCGCGGCCACACACTGCACACAGCTCTCCACCTCTCAGGCGCGAGGAGTTCGTC	1523
Db	1378	GGGAGTGGCGGCCATAGTTTGACCAAGCTCACCCACCTCTGAGGCTGAGGACTTCGTC	1437
QY	1524	TCCGCTCTCCACCCAGAACTACTTCCGTCTCCCTGCCCGGAGCGACCAAGCAATGACC	1583
Db	1438	TCCGCTCTCCACCCAAAATACTTTCGTTCCTGCCCGCGGCAACCAATAGGCC	1497
QY	1584	TATGGGACCTTCAACTTCCTGGGGCGGCGCTGATGATCCCTAATACAGTATCAGCCTC	1643
Db	1498	TACGGGACCTTCAACTTCCTGGGGCGGCGCTGATGATCCCTAATACGGGGATCAGCCTC	1557
QY	1644	CTCATCCCCCAGATGTCATACCCGAGGAAAGATCTATGAGATCTACCTTCACGCTGCAC	1703
Db	1558	CTCATCCCCCGGATGCCATCCCCCGAGGAAAGATCTACGAGATCTACCTTCACATTCGAC	1617
QY	1704	AAGCGGAAGAGTGTAGTTGCCCTTAGTGTGGCTGTACAGCCCTGTGTGAGTCCCATCGTT	1763
Db	1618	AAGCSAGAACAGTGTAGTTGCCCTTAGTGTGGCTGTACAGCCCTGTGTGAGTCCAGTCTG	1677
QY	1764	AGCTGTGGACCCCTCGCGTCTGTCTCACCGCGAGTCACTCTGTGCTATGACACACTGT	1823
Db	1678	AGCTGTGGGCCCCAGAGTCTGTGCTCACCGGCGAGTCACTCTGCAATGGACCACTGT	1737
QY	1824	GGGAGCCCCAGCCCTGACAGCTGGAGCCTTCGCGCTCAAAAAGCAGTCTGTGGAGGCGAGC	1883
Db	1738	GGAGAGCCAGCCCTGACAGCTGGAGTCTGCGCTCAAAAAGAGATCTTCGAGGGCGAGT	1797
QY	1884	TGGGAGAGGATGTGTGTGCACTGTGGCGAGGAGGCGCCTCCCACTCTACTACTGCCAG	1943
Db	1798	TGGG---AGGATGTGTGCACCTTGTGTAGGAGTCACTTCCCACTCTACTACTGCCAG	1854
QY	1944	CTGGAGCCAGTCCCTGCTACGCTCTTCACGAGCAGCTGGGCGGCTTTCCTCTGTGTGGA	2003
Db	1855	CTGGAGCCGGGCGCTGTCTATGTCTTCAGGAGCAGCTGGGCGGCTTTGCCCTGTAGGA	1914

Db 247 GAGGTGCCCCCTAGTGGCTGTCACACCTGCTGAGTCCATCGTTAGTGTGGACCC 306
QY 1778 TGGCCCTCTGCTCACCAGCCAGTATCTCTGCTATGACCACTGTGGGAGCCAGCC 1837
Db 307 TGGCGTCTGCTCACCAGCCAGTATCTCTGCTATGACCACTGTGGGAGCCAGCC 366
QY 1838 TGACAGCTGGAGCTGGCTCAAAAGAGCTGTCGAGGCGAGCTGGGAGCAGATGT 1897
Db 367 TGACAGCTGGAGCTGGCTCAAAAGAGCTGTCGAGGCGAGCTGGGAGCAGTGG--AGATGT 423
QY 1898 GGTGACCTGGGAGGAGGCGCTCCACCTCTACTCTGCTGAGTGGAGCCAGTGC 1957
Db 424 GGTGACCTGGGAGGAGGCGCTCCACCTCTACTCTGAGTGGAGCCAGTGC 483
QY 1958 CTGCTACGCTTTCACGAGCAGTGGGCGCTTTGCCCTGGTGGAGAGCCCTCAGCT 2017
Db 484 CTGCTACGCTTTCACGAGCAGTGGGCGCTTTGCCCTGGTGGAGAGCCCTCAGCT 543
QY 2018 GGCTGCGCAAGCGCTCAAGCTGCTTCTGTTGGGCGGTGGCTGCACTCCCTCGA 2077
Db 544 GGCTGCGCAAGCGCTCAAGCTGCTTCTGTTGGGCGGTGGCTGCACTCCCTCGA 603
QY 2078 GTACAACTCCGGCTCTACTGCTGATGACACCCACGATGCACTCAAGAGGTGTGA 2137
Db 604 GTACAACTCCGGCTCTACTGCTGATGACACCCACGATGCACTCAAGAGGTGTGA 663
QY 2138 GCTGGAGAACGAGCTGGGAGGAGCTGATCCAGGAGCCACGGGTCTGCACTTCAAGA 2197
Db 664 GCTGGAGAACGAGCTGGGAGGAGCTGATCCAGGAGCCACGGGTCTGCACTTCAAGA 723
QY 2198 CAGTTACCAAGCTCGGCTATCCATCCACGATGTCACGATGTCGCTGGAAGATA 2257
Db 724 CAGTTACCAAGCTCGGCTATCCATCCACGATGTCGCTGGAAGATA 783
QY 2258 GCTCCTGTGACGTACAGGAGATCCCTTTTATCATCTGGAATGGACGAGCGGTA 2317
Db 784 GCTCCTGTGACGTACAGGAGATCCCTTTTATCATCTGGAATGGACGAGCGGTA 843
QY 2318 CTTGCACTGCACCTTCAACCTGGAGGTGTCAGCCCGAGCACTAGTGAAGCTGCGTCA 2377
Db 844 CTTGCACTGCACCTTCAACCTGGAGGTGTCAGCCCGAGCACTAGTGAAGCTGCGTCA 903
QY 2378 GCTGTGGGTGTCAGGTGAGGCGGACGAGGAGCTTCAAGTCACTCAATCACTCAATCAC 2437
Db 904 GCTGTGGGTGTCAGGTGAGGCGGACGAGGAGCTTCAAGTCACTCAATCACTCAATCAC 963
QY 2438 CAAAGCACAAAGCTTTGCTGAGTGTGCTGCTGAGAGAGTGAAGCGGGGTCCAGCCCT 2497
Db 964 CAAAGCACAAAGCTTTGCTGAGTGTGCTGCTGAGAGTGAAGCGGGGTCCAGCCCT 1023
QY 2498 GGTGGGCGGAGTGCCTTCAAGATGCCCTTTCATTTGGGAGAGATATTTCCAGCT 2557
Db 1024 GGTGGGCGGAGTGCCTTCAAGATGCCCTTTCATTTGGGAGAGATATTTCCAGCT 1083
QY 2558 GGAACCACTGTAGGCGGGTCCGAGTGGCGGCTCTGGCCAGAACTCCACCTGGA 2617
Db 1084 GGAACCACTGTAGGCGGGTCCGAGTGGCGGCTCTGGCCAGAACTCCACCTGGA 1143
QY 2618 CAGCCATCTCAGCTTCTTTCCTTCAAGCCAGCCCGGAGAGTCACTCAACCTGTG 2677
Db 1144 CAGCCATCTCAGCTTCTTTCCTTCAAGCCAGCCCGGAGAGTCACTCAACCTGTG 1203
QY 2678 GAGGCGGCGACTTCCCAAGCGCAACTCAGCAGCTGCTGCTGAGCAGTGGCTGACT 2737
Db 1204 GAGGCGGCGACTTCCCAAGCGCAACTCAGCAGCTGCTGCTGAGCAGTGGCTGACT 1263
QY 2738 GGCAGCAGCAGCTGCGCTCTTTCAGATGTCGAGGCTGAGTGTGAGGC 2789
Db 1264 GGCAGCAGCAGCTGCGCTCTTTCAGATGTCGAGGCTGAGTGTGAGTC 1315

RESULT 10

AAH99530

ID AAH99530 standard; cDNA; 1321 BP.

XX

AC AAH99530;

XX

DT 16-OCT-2001 (first entry)

XX

DE Human protein encoding cDNA sequence SEQ ID NO:365.

XX

Human; cancer; ulcer; HIV infection; human immunodeficiency virus; antiinflammatory; antirheumatic; antiarthritic; immunosuppressive; antibacterial; endocrine; cardiant; central nervous system; virucide; anti-HIV; fungicide; antimutagen; cardiovascular; antianaemic; anaemia; antiagregant; haemostatic; vulnary; antiulcer; osteopathic; eczema; dermatological; antiallergic; antiasthmatic; antidiabetic; cytostatic; neuroprotective; antidepressant; nootropic; antiparkinsonian; infection; immunostimulant; gene therapy; antisense therapy; vaccine; inflammation; antianaphylactic; rheumatoid arthritis; septic shock; pancreatitis; cardiac dysfunction; neuropathology; cardiac anaphylaxis; autoimmunity; genetic disease; haematopoietic disorder; platelet disorder; asthma; thrombocytopaenia; osteoporosis; severe combined immunodeficiency; allergic rhinitis; diabetes; multiple sclerosis; depression; Alzheimer's disease; Parkinson's disease; neurodegenerative disorder; neurological disorder; ss.

Homo sapiens.

OS

XX WO200153455-A2.

PN

PD 26-JUL-2001.

XX

XX 22-DEC-2000; 2000WO-US035017.

PF

XX 23-DEC-1999; 99US-00471275.

XX

PR 21-JAN-2000; 2000US-00488725.

PR

XX 25-APR-2000; 2000US-00552317.

XX

(HYSE-) HYSEQ INC.

PA

XX Tang YT, Liu C, Drmanac RT;

PI

XX WPI; 2001-457603/49.

XX

DR P-PSDB; AAM25589.

XX

Isolated human polynucleotides encoding polypeptides, useful for the treatment and diagnosis of e.g. cancer, ulcers and HIV infection.

PT

Claim 1; Page 471; 1217pp; English.

XX

AAH99166 to AAH99904 encode the human proteins given in AAM25225 to AAM25963. The proteins can have activities based on the tissues and cells they are expressed in, such as: antinflammatory; antirheumatic; antiarthritic; immunosuppressive; anti-HIV; fungicide; endocrine; cardiant; central nervous system; virucide; anti-HIV; fungicide; antimutagen; cardiovascular; antianaemic; antiagregant; haemostatic; vulnary; antiulcer; osteopathic; dermatological; antiallergic; antiasthmatic; antidiabetic; cytostatic; neuroprotective; antidepressant; nootropic; antiparkinsonian; and immunostimulant. The proteins and polynucleotides encoding them can be used in gene therapy, antisense therapy and vaccine production. The proteins and polynucleotides are useful for screening for agonists or antagonists of a protein and for the treatment and diagnosis of disorders associated with the activity of a protein e.g. inflammation, rheumatoid arthritis, septic shock, pancreatitis, cardiac dysfunction, neuropathology, cardiac anaphylaxis, viral, bacterial, HIV and fungal infections, autoimmunity, genetic diseases, haematopoietic disorders, anaemia, platelet disorders, thrombocytopaenia, wounds, burns, ulcers, osteoporosis, severe combined immunodeficiency, eczema, allergic rhinitis, asthma, diabetes, cancer, multiple sclerosis, depression, Alzheimer's disease, Parkinson's disease, neurodegenerative and neurological disorders

Sequence 1321 BP; 258 A; 440 C; 371 G; 252 T; 0 U; 0 Other;

SQ

Query Match		41.4%; Score 1193.6; DB 4; Length 1321;
Best Local Similarity		97.8%; Pred. No. 1.4e-210;
Matches 1295; Conservative		0; Mismatches 19; Indels 10; Gaps 8;
QY	1476	CACACTGACACAGCTCTCCACCTCTGAGCGGAGGTTCTTCGCGCTCTCC 1535
Db	1	CACACTGACACAGCTCTCCACCTCTGAGCGGAGGTTCTTCGCGCTCTCC 60
QY	1536	ACCCAGACTACTTTCGCTCCCTGCCCGAGGACACAGCAATGACCTATGGACCTTC 1595
Db	61	ACCCAGACTACTTTCGCTCCCTGCCCGAGGACACAGCAATGACCTATGGACCTTC 120
QY	1596	AATCTCTCGGGCGCGCTGATGATCCCTAATACAGGTATCAGCTCTCAATCCGCCCA 1655
Db	121	AATCTCTCGGGCGCGCTGATGATCCCTAATACAGGTATCAGCTCTCAATCCGCCCA 180
QY	1656	GATGCATATCCCGAGGAGATCTATGATCTACTCAGCTGACAAAGCGGAGAAC 1715
Db	181	GATGCATATCCCGAGGAGATCTATGATCTACTCAGCTGACAAAGCGGAGAAC 240
QY	1716	GTGAGTTTCCCTAGCTGGCTGTGAGCCCTGCTGAGTCCCATCGTTAGCTGTGACCC 1775
Db	241	GTGAGTTTCCCTAGCTGGCTGTGAGCCCTGCTGAGTCCCATCGTTAGCTGTGACCC 300
QY	1776	CCT-GGCGTCTCTCAACCGGCCAGTCTATCCT-GGCTATGACCACTGT-GGGGAGCCC 1832
Db	301	CCTGGCGTCTCTCAACCGGCCAGTCTATCCT-GGCTATGACCACTGT-GGGGAGCCC 360
QY	1833	AGCCCTGACAGT-GGAGCTTGGCTCAAAAGCAGTCTGCGAGGGCAGTGGAGCA 1891
Db	361	AGCCCTGACAGTGGAGCTTGGCTCAAAAGCAGTCTGCGAGGGCAGTGGG--A 417
QY	1892	GGATGTCTGACCTGGGAGGAGGCGCTCCACCTCTACTACTGCTCCAGCTGGAGC 1951
Db	418	GGATGTCTGACCTGGGAGGAGGCGCTCCACCTCTACTACTGCTCCAGCTGGAGC 477
QY	1952	CAGTGCCTGCTACGTCTTACCGAGCAGCTGGGCGCTTTGGCTGGTGGAGAGCCCT 2011
Db	478	CAGTGCCTGCTACGTCTTACCGAGCAGCTGGGCGCTTTGGCTGGTGGAGAGCCCT 537
QY	2012	CAGCTGGTGGCGGCAAGCGCTCAAGCTGCTTCTGTTTGGCGGCTGGCTGACCTC 2071
Db	538	CAGCTGGTGGCGGCAAGCGCTCAAGCTGCTTCTGTTTGGCGGCTGGCTGACCTC 597
QY	2072	CCTCAGTACACATCCGGCTTACTGCTGATGACACCCAGCTGACCTCAAGAGGT 2131
Db	598	CCTCAGTACACATCTGCTTACTGCTGATGACCTCAGTATGACCTCAAGTGT 657
QY	2132	GGTGCAGTGGAGAGCAGCTGGGGGACAGCTGATCCAGGACCGGCTCTGCATTT 2191
Db	658	GGTGCAGTGGAGAGCAGCTGACGGGACAGCTGATCCAGGACCGCTGATCTGCACTT 717
QY	2192	CAAGGACAGTTACCAACACCTGGGCTTATCCATCCAGATGGCCAGCTCTGTGAA 2251
Db	718	CAAGGACAGTTACCAACACCTGGGCTTATCCATCCAGATGGCCAGCTCTGTGAA 777
QY	2252	GAGTAAGCTCTTGTGCTACAGAGATCCCTTTTATCACATCTGGAATGGCACCA 2311
Db	778	GAGTAAGCTCTTGTGCTACAGAGATCCCTTTTATCACATCTGGAATGGCACCA 837
QY	2312	GCGTATCTGCACTGCACTTCACTCAGCTGATCCAGCCAGCACTAGTATGCTGGC 2371
Db	838	GCGTATCTGCACTGCACTTCACTCAGCTGATCCAGCCAGCACTAGTATGCTGGC 897
QY	2372	CTGCAGCTGGTGGTGGAGTGGAGGCAAGGAGGAGTTCAGATCAACTTCAA 2431
Db	898	CTGCAGCTGGTGGTGGAGTGGAGGAGGAGTTCAGATCAACTTCAA 957
QY	2432	CATCACCAGGACACAAGGTTTGTGAGCTGCTGGCTGGAGTGAAGGGGGTCC 2491
Db	958	CATCACCAGGACACAAGGTTTGTGAGCTGCTGGCTGGAGTGAAGGGGGTCC 1017
QY	2492	AGCCCTGGTGGGCCCCAGTGGCTTCAAGATCCCTTCTCATTGGCAGAGATAATTC 2551
Db		

Db	1018	AGCCCTGGTGGGCCCCAGTGGCTTCAAGATCCCTTCCCTCATTGGCAGAGATAATTC 1077
QY	2552	CAGCTTGGACCCACCCCTGTAGCGGGGTGCCGACTTGGCGGACTCTGGCCAGAACTCCA 2611
Db	1078	CAGCTTGGACCCACCCCTGTAGCGGGGTGCCGACTTGGCGGACTCTGGCCAGAACTCCA 1137
QY	2612	CTTGGACAGCCATCTCAGCTTCTTTTGGCTTCCAAAGCCAGCCACAGCCATGATCTCAA 2671
Db	1138	CTTGGACAGCCATCTCAGCTTCTTTTGGCTTCCAAAGCCAGCCACAGCCATGATCTCAA 1197
QY	2672	CTTGGTGGAGGGCGGCACTTCCCAACGGCACTTCCCAAGCGCACTCAGCCAGCTGGCTGAGG 2731
Db	1198	CTTGGTGGAGGGCGGCACTTCCCAAGCGCACTTCCCAAGCGCACTCAGCCAGCTGGCTGAGG 1257
QY	2732	T-GGACTGGGCCAGCAGCAGCTGGCTC-TTCACAGTG-TGGAGGCTGAGTGTCTGAGG 2788
Db	1258	TGGACTGGGCCAGCAGCAGCTGGCTC-TTCACAGTG-TGGAGGCTGAGTGTCTGAGG 1317
QY	2789	CCGG 2792
Db	1318	CCGG 1321

RESULT 11
AAS75738
ID AAS75738 standard; cDNA; 3646 BP.
XX
AC AAS75738;
XX
DT 13-FEB-2002 (first entry)
XX
DE DNA encoding novel human diagnostic protein #11542.
XX
KW Human; chromosome mapping; gene mapping; gene therapy; forensic;
KW food supplement; medical imaging; diagnostic; genetic disorder; ss.
XX
OS Homo sapiens.
XX
PN WO200175067-A2.
XX
PD 11-OCT-2001.
XX
PF 30-MAR-2001; 2001WO-US008631.
XX
PR 31-MAR-2000; 2000US-00540217.
PR 23-AUG-2000; 2000US-00649167.
XX
PA (HYSE-) HYSEQ INC.
XX
PI Drmanac RT, Liu C, Tang YT;
XX
DR WPI; 2001-639362/73.
DR P-PSDB; ABG11551.
XX
PT New isolated polynucleotide and encoded polypeptides, useful in
PT diagnostics, forensics, gene mapping, identification of mutations
PT responsible for genetic disorders or other traits and to assess
PT biodiversity.
XX
PS Claim 1; SEQ ID NO 11542; 103pp; English.
XX
CC The invention relates to isolated polynucleotide (I) and polypeptide (II)
CC sequences. (I) is useful as hybridisation probes, polymerase chain
CC reaction (PCR) primers, oligomers, and for chromosome and gene mapping,
CC and in recombinant production of (II). The polynucleotides are also used
CC in diagnostics as expressed sequence tags for identifying expressed
CC genes. (I) is useful in gene therapy techniques to restore normal
CC activity of (II) or to treat disease states involving (II). (II) is
CC useful for generating antibodies against it, detecting or quantitating a
CC polypeptide in tissue, as molecular weight markers and as a food
CC supplement. (II) and its binding partners are useful in medical imaging
CC of sites expressing (II). (I) and (II) are useful for treating disorders

CC involving aberrant protein expression or biological activity. The
CC polypeptide and polynucleotide sequences have applications in
CC diagnostics, forensics, gene mapping, identification of mutations
CC responsible for genetic disorders or other traits to assess biodiversity
CC and to produce other types of data and products dependent on DNA and
CC amino acid sequences. AAS64197-AAS94564 represent novel human diagnostic
CC coding sequences of the invention. Note: The sequence data for this
CC patent did not appear in the printed specification, but was obtained in
CC electronic format directly from WIPO at
CC ftp.wipo.int/pub/published_pct_sequences
XX
SQ Sequence 3646 BP: 930 A; 917 C; 921 G; 877 T; 0 U; 1 Other;

Query Match	31.8%	Score 916	DB 5	Length 3646	
Best Local Similarity	60.3%	Pred. No. 2.1e-159			
Matches 1657	Conservative	1	Mismatches 1016	Indels 72	Gaps 6
Qy	98	GCCCGGCTGTGGCAGCGCTCTGGGCATAGTCTCGCGCTTGGCTCGCGGCTCGGG	157		
Dd	220	GCTGTGCTACCTGCGCTGTCTCAGCGCCAGCGGCACTGGTCCGCCCCCAAGA	279		
Qy	158	TGCCAGCAGAGTGCCACCGTGGCCAAACCAGTGGCTGTGTGCCAAACCGGACCTGTTCC	217		
Dd	280	TGATGACTTTTTTCATGAATCCCCAGAAAATTTTCTTCTGTATCCACTGAGCCTCTGCC	339		
Qy	218	CCACTTCTGGTGAGCGCCGAGGATGTACATCGTCAAGAACCAAGCCAGTGTCTGT	277		
Dd	340	ACATTTCTTATTGAGCCTGAAGAAGCTTATTTGTGAAGAAATAGCCTGTGAACCTGTA	399		
Qy	278	GTGCAAGGCGGTGCCGCCACCGCAGATCTTCTTCAAGTGCACAGGGAGTGGGTGGCCA	337		
Dd	400	CTGTAAAGCAGCCCTGCCACCAGATCTATTTCAAGTGTAAATAGTGAATGGGTTCATCA	459		
Qy	338	GGTGCACCATGTGATCGAGCGCACACAGAGGGGAGCAGTGGTGAGCCGACATGAGGAT	397		
Dd	460	GAAGGACCATAGTAGATGAAGAGTAGATGAACCTTCCGGTCTCATTTCTCGGGAAGT	519		
Qy	398	CCGATTAAATGCTCTCAAGGCAGCAGGTTCGAGAAGGTGTTCGGGTGGAGGAATACTGGTG	457		
Dd	520	GAGCATTCAGATTTCCGCCCAGCAAGTGGAGAACTCTTTGGACCTGAAGATTACTGGTG	579		
Qy	458	CCAGTGGGTGCATGAGACTCTCTCGGGACACCAAGAGTCAGAGGCTTACATCCGCAT	517		
Dd	580	CCAGTGTGTGCGCTTGGAGCTCCGCGGCTACCCACAAAGCGGAAGCGTATGTGGCGCAT	639		
Qy	518	AGCCAGATTGCGAAGAACTTCGAGCAGGAGCGCTCGCCAAAGAGGTGTCCTCGAGCA	577		
Dd	640	TGCATATCTACGGAAGACATTTGAGCAGGAACCCCTAGGAAGGAAGTGTCTTTGNACA	699		
Qy	578	GGGCATCGTGTGCCCTCGCGTCCACCGGAGGCGATCCCTCCAGCCGAGTGGAGTGGCT	637		
Dd	700	GGAAGTCTTACTCCAGTGTGCACCACCTGAAGGGATCCAGTGGCTGAGGTGGAATGGTT	759		
Qy	638	CCGGAACGAGGACCTGTGGACCGCTCCCTGGACCCCAATGTATACATCAGCGGAGCA	697		
Dd	760	GAAAAATGAAGACATAATTGATCCCGTTGAAGATCGGAATTTTATATTACTATTATCA	819		
Qy	698	CAGCCTGTGTGTCGACAGGCGCCGCTTGTGTACACGCGCAACTACACCTGCGTGGCCAA	757		
Dd	820	CAACCTCATATAAGCAGGCGCCGACTCTCTGACTGTGCAAAATACACCTGTGTGCCAA	879		
Qy	758	GAAATGTGTGCACGTGCGCGCAGGCGCTCGCGTGTGTGTATCTCTACGTGAACGCTGG	817		
Dd	880	AAACATTGTGCCAAGAGGAAGAATGACAACCTGCCACTGTCTATGTCTATGTCAACGCTGG	939		
Qy	818	GTGTCGACGTGGACCGAGTGGTCCGCTCTGCAGCGCCAGCTGTGGCGCGCGCTGGCAGAA	877		
Dd	940	CTGGTTCACCTGGACCGAGTGGTCTGTGTGTAAACAGCCGCTGTGCAAGAGGATCAGAA	999		
Qy	878	ACGGAGCCGAGCTGTGCACCAACCCGCGCGCTCTCAACCGGGGGCGCTTCTGTGAGGGCA	937		
Dd	1000	ACGTACAAGACTTGTACAACCCGGGACCACTCAATGGGGGTGCCCTTCTGTGAAGGCA	1059		

Qy	938	GAATGTCATGACCGCACCGTCTCCTCTCTGCTGTGCTCTGTGGACGGCAGCTGGAGCCC	997
Db	1060	GAGTGTGCGAGAA--AATAGCCTGTACTACTGTTATGCCAGTGGATGGCAGTGGACGCC	1116
Qy	998	GTGAGCAAGTGTGGCGCTGTGGCTGGACTGCACCCACTGGCGAGCGCTGAGTGCCTC	1057
Db	1117	ATGGAGCAAGTGTCTACTTGTGGAACCTGAGTGCACCCACTGSCGCGAGGAGGTGCAC	1176
Qy	1058	TGACCCAGCAACCCGCAACGGAGGGGAGGAGTGCCACAGGGGACTGACTGGACACCCGGAA	1117
Db	1177	GGCGCCAGCCCCAAGAAATGGAGCGCAAGACTCGAGCGCCTCGTTTGCATTCCAATCCAGAA	1236
Qy	1118	CTGTACCAGTGAACCTCTGTGTACACAGTGCCTCTGGCCCTGAGGAGTGGCCCTCTATGT	1177
Db	1237	CTGCATGTATGGGCTTTCATATGACAGACTGCTCTGATTAGATGATGTTGCTCTCTATGT	1296
Qy	1178	GGG---CCTCATGCGCGTGCGCTGTGCTGTGCTGCTGCTGCTGCTCTCATCTCGT	1234
Db	1297	TGGGATTTGATAGCAGTGAATCGTTTGGCTGGCGATCTCTGTAGTTGTGGCCTGTGTTGT	1356
Qy	1235	TTATTGCCGAAGAGGAGGGCTTGACATCAGATGTGGCTGACTGCTCCANTCTCACCTC	1294
Db	1357	GTATCGGAAGAATCATCGTCACTTTGAGTCAGATATTTATTGACTCTTGGCATTCAATGS	1416
Qy	1295	AGGCTTTCAGCCCGTCAAGCATTAAGGCCACCAAGCAGACAAACCCCATCTGTCTACCAT	1354
Db	1417	GGGCTTTCARCTGTGAACATCAAG-----GCAGCAAGACAAGATCTGTGGCTGT	1467
Qy	1355	CCAGCGGAGCTCAGACACACACACACCTACAGGGCAGTCTCTGTCCCCGGCAGAGTGG	1414
Db	1468	ACCCCCAGACCTCACGTCACTCGAGCCATGTATCAGAGGACCTGTCTATGCCCTGGATGA	1527
Qy	1415	GCCAGCCCCAAGTTCCAGCTCAACAAATGGGCACTGCTCAGCGCCCCGTGGGTGGCGGCG	1474
Db	1528	CGTCTCAGACAAATCCAAATGACCAACTCTCCAAATCTGGATCCACTGCCCAACCTGAA	1587
Qy	1475	CCACACTGTGACACAA-----GCTCTCCACCTCTGAGGCCGAGAGTTCGT	1522
Db	1588	AATCAAAGTGTACAACACCTCAGTGTGTCTCCCCCAAGATGACCTCTCAGTATTAC	1647
Qy	1523	CTCCCGCTCTCCACCAGAACTACTCCGCTCCCTGCCCGGAGGCACAGCAACATGA-	1581
Db	1648	GTCAAAGTGTCCCTCAGATGACCCAGTCGTTGTTGGAGAAATGAAGCCCTCAGCCTGAA	1707
Qy	1582	-----CCTATGGGACCTTCAACTT	1600
Db	1708	GAACCAGTCTAGCAAGGCAGACTGATCCATCTGTACCGCAATTTGGCAGTTCACATC	1767
Qy	1601	CCTCGGGGCGGTGATGATCCCTAATAAGGTATCAGCTCAGCTCTCATCCCCCAGATGC	1660
Db	1768	GCTGGGAGGTCACTTATTTGTTCCAAATCAGGAGTCAGCTTGTCTGATTCGCCCTGGGGC	1827
Qy	1661	CATACCCCGAGGGAAGTCTATGAGATCTACCTCAGCTGCACAAAGCCGGAAGACGTGAG	1720
Db	1828	CAITCCCCAAGGAGAGTCTACGAATGTTATGTACTGTACACAGGAAGAAACAATATGAG	1887
Qy	1721	GTTGGCCCTAGCTGGCTGTGACACCTCTGAGTCCCATCTGTTAGCTGTGGAACCCCTGG	1780
Db	1888	GCCACCATGATGACTCTCAGACACTTTTGAACCCCTGTGGTAGCTGTGGGCCCCCAGG	1947
Qy	1781	CGTCTGTCTACCCGGCGAGTCACTCTGGCTATGGACCACTGTGGGAGGCCACAGCCCTGA	1840
Db	1948	AGCTCTGCTACCCGCGCGCTGCTCTCATATGATCATCTGCGCAGACCCCAATACCGA	2007
Qy	1841	CAGCTGGAGCTGCGCTCTCAAAAAGCAGTGTGCGAGGGCAGCTGGGAGCAGGATGTGCT	1900
Db	2008	GGACTGGAATAATACTGTCTAAGAACACAGGAGCAGCAGGAGTGGG---AGGATGTGGT	2064
Qy	1901	GCACCTGGCGAGGAGGCGCCCTCCACCTCTACTCTGCGCAGCTGGAGGCCAGTGCCTG	1960
Db	2065	GGTGTGCGGAGGAGAAATTCACACCCCTGCTACATTAAGCTGATGACAGAGGCTG	2124
Qy	1961	CTAGCTCTTCCACGAGCAGCTGGGCGCGCTTTGCCCTGTGGGAGAGGCCCTCAGCGTGGC	2020

DB 2125 CCACATCTCTCACAGAGAACTCAGCACCTAGCCCTCTGGTAGGACATTCACACCAAGC 2184
QY 2021 TCGCGCCAGCGCTCAAGCTCTTCTGTTTGGCGCGTGGCTGACCTCTCCCTCGAGTA 2080
DB 2185 GGTGCAAGCGCTCAAGCTGGCCATCTTTGGGCGCTGTGCTCTCTCGCTGAGTA 2244
QY 2081 CAACATCCGGTCTACTGCTGCTGATCAGACACCCAGATGCTCAAGAGGTGGTGCAGCT 2140
DB 2245 CAGCATCCGAGTCTACTGCTGATGACACCCAGGATGCCCTGAGGAAATTTTACATCT 2304
QY 2141 GGAGAGAGCTGGGGGAGCAGCTGATCCAGAGCCACCGGCTCTGCTCACTTCAAGGACAG 2200
DB 2305 TGAGAGACAGCGGAGGAGCAGCTCTAGAGAACTTAAGGCTCTTCAATTTAAAGGAG 2364
QY 2201 TTACACACACCTCGGCTCTATCCATCCAGATGTGCCAGCTCTCTGTGGAAGTAACT 2260
DB 2365 CACCCACACCTCGGCTCTGCAATTCAGATATCGCCCATTCCTCTGGAAGAGCAAT 2424
QY 2261 CTTGTGAGCTACAGGAGATCCCTTTTATCACATCTGGAATGGCACGCGGTACTT 2320
DB 2425 GCTGCTAAATATCAGGAATTCATTTTACATGTTTGGAGTGGATCTCAAGAACTT 2484
QY 2321 GCATGCACTTCACTCGGAGGTGTAGCCCGCAGCACTAGTGAAGCTGGCTGCAAGCT 2380
DB 2485 GCATGCACTTCACTCTGGAAGATTTAGCTGAACACAGTGGAGCTGGTTTGCAACT 2544
QY 2381 GTGGGTGTGGAGGTGGAGGCGAGCGGAGAGCTTCAGATCACTCAATCAACCA 2440
DB 2545 CTGTGTGGCGAGGTGGAGGAGAGGCGAGATCTTCCAGCTCAACTGCACCGTGTGAG 2604
QY 2441 GGACACAGGTTTGTGAGCTGCTGGCTCTGAGAGTGAAGCGGGTCCAGCGCTGT 2500
DB 2605 GGAACCTACTGSCATCGATTTGCCCTGCTGATCTCTGGAAACACATCACACGCTAC 2664
QY 2501 GGGCCCCAGTCTTCAAGATCCCTTCTCTATTCGGCAGAGATAATTCAGCGCTGGA 2560
DB 2665 GGGCCCCAGTCTTCAAGATCCCTTCTCTATTCGGCAGAGATAATTCAGCGCTGGA 2724
QY 2561 CCCACCTGTAGGCGGTGGAGTGGAGTGGAGTGGAGTGGAGTGGAGTGGAGTGGAG 2620
DB 2725 TGCCCCCAGAGAGAGGCGATGACTGGAGGATGCTGGCCATAGCTGAACCTGGAG 2784
QY 2621 CCATCTGAGCTTCTTGGCTTCAAGCCAGCCAGCCAGCCAGCTATCTCAACCTGGGA 2680
DB 2785 GTACTTGAATTTACTTTGCCAATTCAGCCCACTGGGTAATCTTGGATCTTTGGGA 2844
QY 2681 GGGCGGCACTTCCCAAGCGGAACTCAGCAGCTGGCTGAGCAGTGGCTGGAGTGGG 2740
DB 2845 AGCAGAGAACTTCCAGATGAAACCTGAGCATGCTGGCAGCTGTCTTGGAGAAATGG 2904
QY 2741 CCAGCGAGAGCTGGCTCTTCAAGTGTGGAGGCTGAGTGGTA 2786
DB 2905 AAGACATGAACGGTGTGCTCTTAGCAGCAGAGAGGGCAGTATTAA 2950

RESULT 12
ABT06279
ID ABT06279 standard; cDNA; 2860 BP.
XX
AC ABT06279;
XX
DT 24-OCT-2002 (first entry)
XX
DE Human NOV1a coding sequence.
XX
KW Human; NOVX; autoimmune disease; cancer; infection; inflammatory disease;
KW storage disorder; muscle disorder; neurodegenerative disorder; neuropathic;
KW developmental defect; neuroprotective; antiparkinsonian; hypotensive;
KW hypertensive; haemostatic; cardiac; angiogenic; dermatological;
KW immunosuppressive; antineoplastic; antitumor; antitumor; anti-HIV;
KW antiparasitic; antiallergic; antitumor; antitumor; antitumor; antitumor;
KW vulnary; anorectic; antidiabetic; immunomodulator; antipsoriatic;

KW nephrotropic; kerolytic; antiulcer; cerebroprotective; anticonvulsant;
KW antinfertility; antimanic; antidepressant; metabolic; cytostatic;
XX tranquilizer; analgesic; gene; ss.
OS Homo sapiens.
XX WO200257450-A2.
PN 25-JUL-2002.
XX
XX 29-NOV-2001; 2001WO-US048922.
XX
XX 29-NOV-2000; 2000US-0253834P.
PR 30-NOV-2000; 2000US-0250926P.
PR 25-JAN-2001; 2001US-0264180P.
PR 20-AUG-2001; 2001US-0313656P.
PR 05-OCT-2001; 2001US-0327456P.
PR 28-NOV-2001; 2001US-00327456.
XX (CURA-) CURAGEN CORP.
PA
XX Edinger S, Macdougall JR, Millet I, Ellerman K, Stone DJ;
PI Gerlach V, Grosse WM, Alsbrogk JP, Lepley DM, Rieser D, Burgess CE;
PI Casman SJ, Spytek KA, Boldog FL, Li L, Fadigaru M, Mishra V;
PI Patturajan M, Shenoy S, Rastelli L, Tchernev VT, Vernet CAM;
PI Zerhusen BD, Malyankar UM, Guo X, Miller CE, Gangolli EA;
XX WPI; 2002-590741/63.
DR P-PSDB; AAO18734.
XX
PT Novel isolated polypeptide, designated NOVX, useful for treating or
PT preventing in NOVX-associated disorders e.g. cardiomyopathy,
PT atherosclerosis, diabetes, cancer, allergy, asthma, Crohn's disease.
XX
PS Claim 9; Page 11; 353pp; English.
XX
CC The present invention provides the protein and coding sequences of
CC several novel human proteins, designated NOVX. These can be used in the
CC treatment of, amongst others, cancers, autoimmune diseases, infections,
CC inflammatory diseases, storage disorders, muscle disorders,
CC neurodegenerative diseases and developmental defects. The present
CC sequence is a coding sequence of the invention
SQ Sequence 2860 BP; 555 A; 950 C; 860 G; 495 T; 0 U; 0 Other;
Query Match 30.9%; Score 889; DB 6; Length 2860;
Best Local Similarity 60.7%; Pred. No. 1.9e-154;
Matches 1637; Conservative 0; Mismatches 960; Indels 102; Gaps 7;
QY 184 ACCAGTGCCTGGTGCCAAACCCGACCTGCTTCCCACTTCTGTGGAGCCGAGGATG 243
DB 168 ACTCTTCCCTGTCAGCGCCAGCAGCGCTGCTCTTCTCTCAGGAGCCACAGGAGC 227
QY 244 TGTACATCGTCAAGAACAGCCAGTGTCTGTGTGTCAGAGCGCTGCTCCCGCAGCAGA 303
DB 228 CCTACATGTGAAGAACAGCCCTGTGAGCTTCCGCTGCGCGCTTCCCGCCACACAGA 287
QY 304 TCTTCTTCAAGTGAACAGCGGAGTGGGTGCGCCAGTGGACCGTGTATCGAGCGCAGCA 363
DB 288 TCTACTTCAAGTGAACAGCGGAGTGGGTGCGCCAGAGCAGCAGTGTACACAGGAAGCC 347
QY 364 CAGAGCGAGAGTGGTGAGCCGACCATGAGGCTCCGATTAATGTCTCAAGCGCAGG 423
DB 348 TGGATGAGGCCACCGGTCTGCGGTGCGGAGGTGAGATCGAGGTGTCTCGCGCAGCAG 407
QY 424 TCGAGAGGTGTTCGGCTGGAGAACTGTGTGCGAGTGGGTGCGATGAGCTCTCGG 483
DB 408 TGGAGAGCTCTTTGGCTGGAGATTAATGTGTGCGAGTGGGTGCGCTGAGCTCCGCG 467
QY 484 GCACCAACCAAGAGTCAGAGGCTTACATCCGATAGCAGATTGCGCAAGACTTCCAGC 543
DB 468 GCACCAACCAAGAGTCGCGGAGCTTACGCTCGCATCGCTTACCTGCGCAAGACTTCGATC 527

Db	2692	TCTCTATGGACGGGTA	CTGTAATTA	CTTTTGGCCACAAAGCAGAGCCCCACAGGGTGTGATCC	2741
Qy	2668	TCAACCTTGGGAGCGCGG	CACCTTCCC	AACGGCAACCTCAGCCAGCTGGCTGCAGCAG	2727
Db	2742	TGGACCTCTGGGAAGCT	CTGCAGCAGACGATGGG	AGCCTCAACAGCTTGGCGAGTGCCT	2801
Qy	2728	TGGCTGGACTGGCGCAG	CCAGCAGCGCTGGC	CTTCTTCAACAGTGTGCGAGGCTGAGTGTGA	2786
Db	2802	TGGAGGAGATGGGCAAG	AGTAGATGCTGTGGT	GTGTGGCCACCGACGGGACTGCTGA	2860
RESULT 13					
ABT06280					
ID	ABT06280	standard;	cdNA;	2860 BP.	
XX	AC	ABT06280;			
XX	DT	24-OCT-2002	(first entry)		
XX	DE	Human NOV1b	coding sequence.		
XX	Human;	NOVX;	autoimmune disease;	cancer;	infection;
KW	storage disorder;	muscle disorder;	neurodegenerative disorder;	inflammatory disease;	
KW	developmental defect;	neuroprotective;	antiparkinsonian;	hypotensive;	
KW	hypertensive;	haemostatic;	cardiant;	antiangiinal;	dermatological;
KW	immunosuppressive;	antiinflammatory;	virucide;	antibacterial;	anti-HIV;
KW	antiparasitic;	antiallergic;	antiasthmatic;	immunomodulator;	antipsoriatic;
KW	vulnery;	anorectic;	antidiabetic;	antitumor;	cerebroprotective;
KW	nephrotropic;	kerolytic;	antulcer;	immunomodulator;	anticonvulsant;
KW	antinfertility;	antimanic;	antidepressant;	metabolic;	cytostatic;
XX	tranquillizer;	analgesic;	gene;	ss.	
OS	Homo sapiens.				
XX	WO200257450-A2.				
XX	25-JUL-2002.				
XX	29-NOV-2001;	2001WO-US048922.			
PR	29-NOV-2000;	2000US-0253834P.			
PR	30-NOV-2000;	2000US-0250926P.			
PR	25-JAN-2001;	2001US-0264180P.			
PR	20-AUG-2001;	2001US-0313656P.			
PR	05-OCT-2001;	2001US-0327456P.			
PR	28-NOV-2001;	2001US-00327456.			
XX	(CURA-)	CURAGEN CORP.			
XX	Edinger S,	MacDougall JR,	Millet I,	Ellerman K,	Stone DJ;
PI	Gerlach V,	Grosse WM,	Alsbrook JP,	Lepley DM,	Rieger D,
PI	Casman SJ,	Spytek KA,	Boldog FL,	Li L,	Padigaru M,
PI	Patturajan M,	Shenoy S,	Rastelli L,	Tchernev VT,	Vernet CAM;
PI	Zerhusen BD,	Malyankar UM,	Guo X,	Miller CE,	Gangolli EA;
XX	WPI;	2002-590741/63.			
DR	P-PSDB;	AAO18735.			
DR	Novel isolated polypeptide,	designated NOVX,	useful for treating or		
PT	preventing in NOVX-associated disorders	e.g. cardiomyopathy,			
PT	atherosclerosis, diabetes, cancer,	allergy, asthma, Crohn's disease.			
XX	Claim 9;	Page 14;	353pp;	English.	
XX	The present invention provides the protein and coding sequences of				
CC	several novel human proteins, designated NOVX. These can be used in the				
CC	treatment of, amongst others, cancers, autoimmune diseases, infections,				
CC	inflammatory diseases, storage disorders, muscle disorders,				
CC	neurodegenerative diseases and developmental defects. The present				
CC	sequence is a coding sequence of the invention				
XX	Sequence 2860 BP;	557 A;	950 C;	858 G;	495 T;
SQ	0 U;	0 Other;			

Db 1185 TGGCATCTCATGCGGTGGGGTGGTGTGTACCGCGCACTGCGGTGACTTCGACA 1244
Qy 1264 CAGATGTGGCTGACTCGTCCATT---CTACCTCAGGCTTCAGCCCGTACGATCAAGC 1320
Db 1245 CAGACATCACTGACTCATCTGCTGCCCTGACTGGTGTTCACCCCGTCAACTTTAAGA 1304
Qy 1321 CCAGCAAGCAGACAAACCCCATCTGCT-----CACCATCAGCGGACCTCAGACCA 1374
Db 1305 CGCAAGGCCAGTAACCCCGCAGCTCTACACCCCTCTGTGCTCTCTGACCTGACAGCA 1364
Qy 1375 CCACACACTACAGGCGAGTCTGTGTCGCCGACGATGGGCCAGCCCAAGTTCCAGC 1434
Db 1365 GGGCGGACATPACCGGAGCCCGTGTATGCCCTGCAGGACTCCACCGACAAATCCCCA 1424
Qy 1435 TCACCAATGGGCACTGCTCAGCCCC-----CTGGTGGCGCGCC 1460
Db 1425 TGACCAACTCTCTCTGCTGGACCCCTTACCCAGCCTTAAGGTCAAGGTCTACAGTCCA 1484
Qy 1461 -----CTGGTGGCGCGCC 1476
Db 1485 GCACACGGGCTCTGGGCCAGGCTGGGAGATGGGGCTGACCTGTGGGGTCTTGGCGC 1544
Qy 1477 ACACACTGCACACAGCTCTCCACCTCTGAGCGGAGGAGTTCGTCTCCCGCTCTCCA 1536
Db 1545 CTGGCACATACCTAGGATTTGCGCCGGAACACCACTCTCTGCACTGCGCAGCGCA 1604
Qy 1537 CCCGAACTAC-----TTCCGCTCCCTGCGGAGGACACGACAACTGACTTGGGA 1590
Db 1605 GCCTGGTTCACAGCAGCTCTTGGGCTGCGCCGAGACCCAGGAGCAGCGTCAAGCGCA 1664
Qy 1591 CTTTCACTCTCGGGGCGGCTGATGATCCCTAATACAGGTATCAGCCTCCATCC 1650
Db 1665 CTTTGGCTGCTGGGTGGAGGCTCAGCATCCCGGACAGGAGTCACTGTGCTGGTG 1724
Qy 1651 CCCAGATGCATACCCCGAGGAAGATCTATGAGATCTACCTCAGCTGCGCAAGCCGG 1710
Db 1725 CCAATGGAGCAATCCCGAGGCAAGTTCTACGAGATGATCTACTCATCAAGGCGAG 1784
Qy 1711 AAGACGTGAGGTTCGCCCTAGCTGTGCTGTGACACCTGTGATGCCATCGTTAGTGTG 1770
Db 1785 AAGTACCTCGCGCTTTTCAGAGGAGCCAGACAGTATTGAGCCCTCGGTGACCTGTG 1844
Qy 1771 GACCCCTGGCTGCTGCTCACCGGCGAGTATCCTGGCTATGACCACTGTGGGGAGC 1830
Db 1845 GACCCACAGGCTCTGCTGTGCGCCCGCTATCTCTACCATGCCCCCTGTGCGCAAG 1904
Qy 1831 CCAGCCCTGACAGCTGGAGCTCGGCTCCTCAAAAGCAGTCTGTGCGAGGCGAGCTGGGAGC 1890
Db 1905 TCAGTCCCGTGAAGTCTTTTCAGCTCAAGACCCAGGCCACAGGGCCACTGGGAG- 1963
Qy 1891 AGGATGTGCTGCACTGGGGAGGAGGCGCTCCCACTCTACTACTGCGCACTGGAGG 1950
Db 1964 --CAGGTGTGACCTCTGGATGAGGAGACCTGAACACACCCCTGCTACTGCGACCTGGAGC 2021
Qy 1951 CGAGTGGCTCTAGCTTTACCGACAGCTGGGCGCTTTGCTGCTGGGAGAGGCC 2010
Db 2022 CCAGGGCTGTACATCTCTGCTGGACAGCTGGGACCTTACGTGTTCAOAGGCGAGTCT 2081
Qy 2011 TCAGCGTGGCTGCGCCCAAGCGCTCAAGCTGCTTCTTTTGGCGCGGTGGCTGCACT 2070
Db 2082 ATTCCCGCTCAGCAGTCAAGCGCTCCAGTGTGCGCGCTTTCGCCCCCGCTCTGCACCT 2141
Qy 2071 CCTTCAGTACATCATCCGGGTCTATGCTGCTGATGACACCCAGATGACTCAAGGAGG 2130
Db 2142 CCCTGGAGTACAGCTCCGGGTCTACTGCTGCGGAGACACCGCTGTAGCACTGAAGGAGG 2201
Qy 2131 TGTGTCAGCTGGAGAACAGCTGGGGGACAGCTGATCCAGGAGCCACCGGTCTCTGCACT 2190
Db 2202 TGTGAGCTGGAGCGGAGCTCTGGCGGATCTTGTGGAGGAGCGCAACCGCTAATGT 2261
Qy 2191 TCAAGGACAGTTACCAACCTGCGCTATCCATCCAGATGTGCCAGCTCCCTGTGGA 2250

Db 2262 TCAAGGACAGTTACCAACCTGCGCTCTCCCTCCATGACCTCCCCCATGCCCATTTGA 2321
Qy 2251 AAGTAACTCTCTGCTAGCTACAGGAGATCCCTTTTATCACAATCTGGATGCGACGC 2310
Db 2322 GGAGCAAGCTGCTGGCCAAATACAGGAGATCCCTTTCTATCACAATTTGGATGCGACGC 2381
Qy 2311 AGCGTACTTTCGACTGCACCTTTCACCTTGGAGCGTGTGAGCCCAAGCACTAGTGAACCTGG 2370
Db 2382 AGAAGCCCTCCACTGCACTTTTACCTTGGAGAGGACACAGCTTGGCTTCCACAGAGTCA 2441
Qy 2371 CTTGCAAGCTGTGGTGTGGCAGGTGGAGGGGAGCGGAGAGCTTTCAGCATCACTTCA 2430
Db 2442 CTTGCAAGATCTGCTGGCGCAAGTGAAGGGGAGGGCCAGATATTCAGCTTGCATACA 2501
Qy 2431 ACATCACCAG---GACACAAGTTTGTGAGCTGCTGCTGCTGAGAGTGAAGCGGGG 2487
Db 2502 CTCTGGCAGAGACACTGCTGGCTCCCTGGGACACTCTCTGCTCTGCCCTGGCAGACTG 2561
Qy 2488 TCCCAAGCCTGTGGGGCCCCAGTGCCTTCAAGATCCCTTCCATTCGGCAGAGATAA 2547
Db 2562 TCACCAACCCAGCTGGGACCTTATGCTTCAAGATCCCACTGCTCCATCCGCCAGAGATAT 2621
Qy 2548 TTTTCAGCTGAGACCCACCTGAGCGGGTCCGACTGGCGGACTCTGGCCCAAGAAC 2607
Db 2622 GCAACAGCTAGATCCCCCACTCACGGGCAATGACTGGCGGATTTAGCACAGAGC 2681
Qy 2608 TCCACTGACAGCACTCTCAGCTTCTTGGCTTCAAGCCCAAGCCCAAGCCCAAGCATGATCC 2667
Db 2682 TCTTAAGACCGTACCTGAATTTGCTCCACCAAGCAGAGCCCAAGCGGTGTATCC 2741
Qy 2668 TCAACCTGTGGAGGCGGCGCACTTCCCAAGCGCAACTCAGCCAGCTGGTGTGAGCAG 2727
Db 2742 TGGACCTCTGGGAGCTCTGACGAGGACGATGGGACCTCAACAGCTTGGCGAGTGCCT 2801
Qy 2728 TGCTGAGCTGGGCGCAGCAGCTGGCTCTTACAGTGTGGAGGCTGAGTGTGA 2786
Db 2802 TGGAGAGATGGGCAAGAGTGAAGTGTGGTGTGGCCACCGAGCGGGACTGTCTGA 2860

RESULT 14

ABK92062
ID ABK92062 standard; DNA; 2995 BP.

AC ABK92062;

XX 14-AUG-2002 (first entry)

DT DNA encoding novel UNC5 receptor-like protein.

DE
XX
XX
KW Cancer; metabolic disorder; skin disorder; infectious disease; anorexia;
KW behavioral disorder; valve disease; endocrine disorder; heart disorder;
KW blood disorder; anxiety disorder; brain disorder; inflammatory disorder;
KW neurodegenerative disorder; Alzheimer's disease; Parkinson's disease;
KW immune disorder; haematopoietic disorder; dyslipidaemia;
KW metabolic disturbance; obesity; metabolic syndrome X; wasting disorder;
KW gene therapy; transgenic animal; human; gene; ds.

XX Homo sapiens.

XX WO200229058-A2.

XX 11-APR-2002.

XX 05-OCT-2001; 2001WO-US031248.

XX 05-OCT-2000; 2000US-0238323P.

XX 05-OCT-2000; 2000US-0238325P.

XX 06-OCT-2000; 2000US-0238372P.

XX 06-OCT-2000; 2000US-0238373P.

XX 06-OCT-2000; 2000US-0238379P.

XX 06-OCT-2000; 2000US-0238382P.

XX 06-OCT-2000; 2000US-0238383P.

XX 06-OCT-2000; 2000US-0238384P.


```
Db 1401 AGCAGCGCCGGGATCTACCGGGGACCGGTGTATGCCCTGCAGACTCCACCGACAAT 1460
Qy 1430 CCAGCTACCAATGGGCACTGCTCAGCCCC----- 1460
Db 1461 CCCATGACCAACTCTCTCTGCTGGACCCCTTACCAGCCTTAAGGTCAGGTCTACAG 1520
Qy 1461 ----- 1471
Db 1521 CTCAGACCAACGGGCTCTGGGCCAGGCTGCGAGATGGGGCTGACCTGCTGGGGTCTT 1580
Qy 1472 CGGCACACATGACACAGCTCTCCACCTCTGAGCGGAGGAGTTGCTCTCCCGCT 1531
Db 1581 GCGGCTGGCAATACCTAGGATTTGCCCGGACACCCACTTCTGCACTGGGAG 1640
Qy 1532 CTCACACC-----AGAACTACTTCGGTCCCTGCCCCGAGCAACAGACCTA 1585
Db 1641 CGCAGCGCTCGGTTCCAGCAGCTCTTGGGCGCTGCCCGAGACCAGGAGCAGGCTCAG 1700
Qy 1586 TGGGACCTTCACTTCTCGGGGCGGCTGATGATCCCTAATACAGGTATCAGCTCTCT 1645
Db 1701 CGGACCTTTGGCTCGCTGGTGGGAGGCTCAGCATCCCCCGCAGGAGGTGTCAGCTTGT 1760
Qy 1646 CATCCCCCAGATGATCATACCCGAGGGAAGATCTATGATCTACCTCACGCTGCACAA 1705
Db 1761 GGTGCCAATGAGCCATTCGCCAGGCGAGTTCTACGAGATGTTACTCATCAACA 1820
Qy 1706 GCGGAAGACGTAGGTTGCCCTAGCTGGCTGTACAGCCCTGCTGAGTCCCATGTTAG 1765
Db 1821 GCGAAGATACCTCGCGCTTTCAAGAGGACCCAGACAGTATTGAGCCCTCGGTGAC 1880
Qy 1766 CTGTGACCCCTGGGCTCTGCTACCGGCGAGTCACTCTGGCTATGACCACTGTGG 1825
Db 1881 CTGTGACCCACAGGCTCTGCTGTGCGGCGCGCTACCTCAACATCCGCTGCTGTC 1940
Qy 1826 GGAGCCAGCCCTGACAGCTGAGCGCTGCGCTCAAAAAGCAGTCTGCGAGGCGAGCTG 1885
Db 1941 CGAAGTCAGTCCCGTCACTGATCTTTTCACTCAAGACCCAGGCCACAGGGCCACTG 2000
Qy 1886 GGAGCAGGATGCTGTCACCTGGGCGAGGCGCCCTCCACCTCTACTTGCAGCT 1945
Db 2001 GGAGCAGGAGGTGGTGACCTGGATGAGGAGACCTTGAACACACACCTGCTACTGCCAGCT 2060
Qy 1946 GGAGGCGAGTCTGCTAGCTCTTACCGAGAGCTGGCGGCTTTGCCCTGGTGGAGA 2005
Db 2061 GGAGCCAGGCGCTGTCATCTGCTGACACAGCTGGGACCTACTGTTTACGGGCGA 2120
Qy 2006 GCGCCTCAGCTGGCTGCGGCAAGCGCTCAAGCTGCTTCTGTTTGGCGCGGTGGCCTG 2065
Db 2121 GTCTATTCCGCTCAGCAGTCAAGCGCTCCAGCTGGCGCTCTTCCGCCCGCCCTCTG 2180
Qy 2066 CACTTCCCTCGAGTACAAATCGGGTCTACTGCTGATGACACCCAGCAGATGACCTCAA 2125
Db 2181 CACTTCCCTGGAGTACAGCTCGGGTCTACTGCTGGAGGACACGCTGTAGACTGAA 2240
Qy 2126 GGAGTGGTGCAGCTGGAGAGCAGCTGGGGGACAGCTGATCAGGAGCCACGGTCTCT 2185
Db 2241 GGAGGTGCTGGAGCTGGAGCGGACTCTGGCGGATCTTGGTGAGAGCGCCGAACCGCT 2300
Qy 2186 GCATTTCAAGGACAGTTACCAACCTGCGCCTATCCATCCAGATGTGCCAGCTCCCT 2245
Db 2301 AATGTTCAAGGACAGTTACCAACCTGCGCCTCTCCCTCATGACCTCCCATGCCCCA 2360
Qy 2246 GTGGAAGTAAGCTCTTGTAGCTACAGGAGATCCCTTTATCAGATCTGGATGG 2305
Db 2361 TTGGAGGAGCAAGCTGTGGCCAAATACCAAGAGATCCCTTCTATCATTGGAGTGG 2420
Qy 2306 CACGACGCGTACTTGCAGTCACTTCACTTCCCTGGAGCGTGTGAGCCCGACACTAGTGA 2365
Db 2421 CAGCAGAGAGCCCTTCCACTGCACTTTTCACTTGGAGAGGACAGCTTGGCTCCACAGA 2480
Qy 2366 CCTGGCTGCAAGCTGTGGGTGTGGCAGGTGGAGGGCGACGGGAGAGCTTTCAGCATCAA 2425
Db 2481 GCTCACTCTCAAGATCTGCGTGGCGCAAGTGGAGGGGAGGGCGAGATATTCAGCTGCA 2540
```

```
Qy 2426 CTTCAACATCACCAAGGACA---CAAGGTTTGTGAGCTGTGGCTCTGGAGAGTGAAGC 2482
Db 2541 TACCACCTGGAGAGACACCTGTGGCTCCCTGGACACTCTCTGCTGTGCCCTGGCAG 2600
Qy 2483 GGGGTCCAGACCTGTGTGGCCCCAGTGCCTTCAAGATCCCTTTCCTCATTCGGCAGAA 2542
Db 2601 CACTGTACACCCAGCTGGGACCTTATGCTTCAAGATCCCACTGTCCATCCGCCAGAA 2660
Qy 2543 GATAATTTCCAGCCTGGAGCCCAACCTCTAGGCGGGGTGCCGACTGGCGGACTCTGCCCA 2602
Db 2661 GATATGCAACAGCTTAGATGCCCACTCACGGGGCAATGACTGGCGGATGTTAGCACA 2720
Qy 2603 GAACTTCCACCTGGAGACCCATCTCAGCTTCTTTCCTTCAAGCCAGCCGCCACAGCCAT 2662
Db 2721 GAAGCTCTCTATGACCCGCTTACCTGAATATCTTTCGCAACCAAGCGGCCACCGGTGT 2780
Qy 2663 GATCTTCAACCTGTGGAGCGCGCACTTCCCAAGCCAACTCAGCAGCTGGCTGC 2722
Db 2781 GATCTGACCTCTCGGAAGCTCTGCAGCAGCAGATGGGACCTCAACAGCTGGCGAG 2840
Qy 2723 AGCAGTGGTGGACTGGGCGCAGCAGCTGGCTCTTTCAGATGTGCGAGGCTGAGTG 2782
Db 2841 TGCTTGGAGGAGATGGGCAAGATGAGATGCTGTGGCTGTGGCCACCGCGGAGCTG 2900
Qy 2783 CTGAGGC 2789
Db 2901 CTGAGCC 2907

RESULT 15
ABQ93898
ID ABQ93898 standard; DNA; 2895 BP.
XX ABQ93898;
XX AC
XX 01-NOV-2002 (first entry)
XX Human transmembrane receptor UNC5H2-like NOV11 DNA, SEQ ID NO:37.
DE Human; NOVX; neurological disorder; Alzheimer's disease;
KW Huntington's disease; Parkinson's disease; pain; behavioural disorder;
KW addiction; tuberous sclerosis; cancer; immune disorder; allergy;
KW autoimmune disease; myasthenia gravis; asthma; arthritis; diabetes;
KW thyroiditis; cardiovascular disease; hypertension; reproductive disorder;
KW endometriosis; incontinence; psoriasis; scleroderma; alopecia; ulcer;
KW pancreatitis; cirrhosis; glomerular endotheliosis; bacterial infection;
KW polycystic kidney disease; endocrine disorder; obesity; cardiomyopathy;
KW atherosclerosis; cell signal processing-related disorder;
KW metabolic pathway regulation disorder; cytostatic; neuroprotective;
KW antiinflammatory; immunosuppressive; analgesic; antithrombotic;
KW dermatological; antibacterial; antiarthritic; hepatotropic; neurogenesis;
KW differentiation; proliferation; motility; haematopoiesis; wound healing;
KW angiogenesis; forensic biology; transgenic animal; drug screening;
KW gene therapy; NOV11; transmembrane receptor UNC5H2-like; chromosome 10;
KW gene; ds.
XX Homo sapiens.
XX Key Location/Qualifiers
FH 5'UTR 1..30 /*tag= a
FT 31..2868 /*tag= b
FT CDS /*product= "NOV11"
FT 31..108 /*tag= c
FT sig_peptide 109..2865 /*tag= d
FT mat_peptide /*product= "Mature NOV11 protein"
FT 2869..2895 /*tag= e
FT 3'UTR
XX
```


QY 1228 TCCTCGTTTATGCGGAGAGGAGGGCTGGACTCAGATGTGGTGACTCGTCAATT- 1286
Db 1217 TGGTGGTGTACCGCGCACTCCCGTGAATTGCGACAGACATCACTGACTCATCTGCTG 1276
QY 1287 --CTCACCTCAGGCTTCCAGCCCGTCCAGCATCAAGCCAGCAAGAGAGAGCAACCCCGCATC 1344
Db 1277 CCCTGACTGGTGGTTTCCACCCCGTCACTTTAAGAGCGCAAGGCCAGTAACCCGCGAGC 1336
QY 1345 TGCT-----CACCATPCCAGCCGACCTCAGCACACACACACCTACACAGGGAGTCTCT 1398
Db 1337 TCTTACACCCCTCTGTGCTCTGCTGACCTGACAGCCAGCCCGGCATCTACCGCGAGCCG 1396
QY 1399 GTCCCGGAGAGTGGGCGCCGCCCAAGTTCAGCTCACCAGTGGGACCTGTCCAGCC 1458
Db 1397 TGTATGCTGAGGACTCCCGCAAAATCCCATGACCAACTCTCTCTGCTGGACC 1456
QY 1459 CC----- 1460
Db 1457 CTTTACCAGCCTTAAGGTCAAGTCTACAGCTCCAGACCCAGGGCTCTGGGCCAGGCC 1516
QY 1461 -----CTGGGTGGCGCGGCCACACACTGACACAGCTCTCCCA 1500
Db 1517 TGGCAGATGGGCTGACCTGCTGGGGTCTTGCCGCTTGCCACATACCTTAGCGAFTTCG 1576
QY 1501 CCTCTAGAGCCGAGGAGTTCGCTCCGCTCTCCACCC-----AGAACTACTTCCGCT 1554
Db 1577 CCGGGACACCACTTCTCTGACCTCGCAGCGCCAGCTCTGGTTCGAGAGCTCTTG 1636
QY 1555 CCTGCTCCGAGGACCAAGCAATGACCTATGGGACCTTCAACTTCTTGGGGGCGCGC 1614
Db 1637 GCTGCCCCGAGACCCAGGGAGCAGCTCAGCGCACCTTTGGCTGCTGGTGGGAGGC 1696
QY 1615 TGATGATPCCCTAATACAGGTATCAGCTCTCTATCCCGCCAGATGCCATACCCGAGGGA 1674
Db 1697 TCAGCATCCCGGCACAGGTGTGAGCTTGTGTGCCCCAATGAGAGCAATCCCGAGGGCA 1756
QY 1675 AGATCTATGAGTCTACTCACGCTGCACAAGCCGGAAGAGTGAAGTTGCCCTAGCTG 1734
Db 1757 AGTTCTACGAGATGTATCTACTCATCAAGGCAAGAGTACCTTGCCGCTTTCAGAG 1816
QY 1735 GCTGTCCAGACCTCTGCTGAGTCCCATGTTAGCTGTGAGACCCCTCGGCTCTCTCACCC 1794
Db 1817 GGAACAGACAGTATGAGCCCTCGGTGACCTGTGAGCCACAGGCCCTCTCTGTGCC 1876
QY 1795 GGCAGTCACTCTGGGTATGGAACACTGTGGGAGCCAGCCCTGACAGTGTAGCTGTC 1854
Db 1877 GCGCCGCTCACTCAACATGCCCCACTGTGCCGAAGTCACTGTCGCGTGAAGTCTTTC 1936
QY 1855 GCCTCAAAAGACAGTCTGTGGAGGAGCTGGGAGAGGATGTGCTGCACTGGCGGAGG 1914
Db 1937 AGTCTAAGACCCAGGCCACAGGGCCACTGGG---AGGAGTGGTGACCCCTGGATGAGG 1993
QY 1915 AGGCGCCCTCCACCTCTACTGCTGAGCTGGAGGCCAGTGGCTGTGCTACTCTTACCG 1974
Db 1994 AGACCTGTAACACACCTGTACTGCTGAGCTGGAGCCAGGCCGTGTACATCTGTCTG 2053
QY 1975 AGCAGTGGGCGGCTTGGCCCTGTGGGAGAGGCCCTCAGGTTGGCTGCGCCCAAGGCC 2034
Db 2054 ACCAGTGGGCACTACGTGTTCAAGGGGAGTCTTATTCGCTCAGAGTCAAGCGGC 2113
QY 2035 TCAAGTGTCTTGTGTCGCGGCTGACCTCCCTCGAGTCAAAATCCGGGTCT 2094
Db 2114 TCCAGCTGGCGGTCTTCCGCCCCGCTCTGACCTCCCTGGAGTACAGCTCCGGGTCT 2173
QY 2095 ACTGCTGTGATGACACCCAGATGCACTAAGGAGGTGGTGCAGCTGGAGAGAGCTGG 2154
Db 2174 ACTGCTGGAGGACACGCTGTAGCACTGAAGGAGGTGCTGGAGCTGGAGCGGACTCTGG 2233
QY 2155 GGGGACAGCTGATCCAGGAGCCAGGGTCTGTGCACTTCAAGGACAGTTACCAAACTGC 2214
Db 2234 GCGGATACTTGGTGGAGGCGGAAACCGCTAATGTTCAAGGACAGTTACCAAACTGC 2293
QY 2215 GCCTATCCATCCAGATGTGCCAGCTCCCTGTGGAAGAGTAAAGCTCTCTGTGAGTACC 2274

Db 2294 GCCTCTCCCTCCATGACCTCCCCCATGCCATTGGAGGAGCAAGCTGTGTGGCAATACC 2353
QY 2275 AGGAGATCCCTTTTATCAGATCTGGAATGGCAAGCGGTACTTGGACTGCACCTTCA 2334
Db 2354 AGGAGATCCCTTTTATCAGATTTGGAGTGGCAGCAGAGCCCTTCCACTGCATTTCA 2413
QY 2335 CCCTGGAGCGTGTCAAGCCCAAGCACTAGTGAACCTGCTGCAAGCTGTGGTGTGGCAGG 2394
Db 2414 CCCTGGAGGACAGCTTGGCTCCACAGAGCTCACCTGCAAGATCTGCTGGGCGCAAG 2473
QY 2395 TGGAGGCGAACCGGAGAGCTTTCAGATCAACTTCAATCAATCAACCAAG---GACACAGGT 2451
Db 2474 TGGAGGCGGAGGCGCAGATATTCAGCTGATACCACTCTGGCAGAGACACCTGTGGCT 2533
QY 2452 TTGCTGAGTGTGGCTCTGGAGAGTGAAGCGGGGTCCAGCCCTGTGTGGGCCCCAGTG 2511
Db 2534 CCCTGGACACTCTGTGCTCTGCCCCCTGGCAGCACTGTCAACCCAGCTGGGACCTTATG 2593
QY 2512 CTTTCAAGATCCCTTTCTCTTTCGCGAGAGATAATTTCCAGCTTGGACCCCACTGTA 2571
Db 2594 CTTTCAAGATCCCACTGTCTCCATCCGCCAGAGATATGCAACAGCTTAGATGCCCACT 2653
QY 2572 GCGCGGCTGCCAGCTGGCGGACTCTGGCCCCAGAACTCCACCTGGACAGCCATCTCAGCT 2631
Db 2654 CACGGGGCAATGACTGGCGGATGTTAGCACAGAAAGCTCTCTATGGACCGGTACTGAATT 2713
QY 2632 TCTTTGCTTCCAGCCAGCCCGCCAGCCATGATCTCACTGTGGGAGGCGCGCACT 2691
Db 2714 ACTTTGCCACCAAGGAGCGCCCCACGGGTGTGATCTCTGGACCTCTGGGAGCTCTGAGC 2773
QY 2692 TCCCCAAACGCAACCTCAGCCAGCTGGCTGCGAGCTGGCTGGACTGGGCGAGCCAGAG 2751
Db 2774 AGGACGATGGGACCTTCAACAGCTGCGGAGTGCCCTTGGAGGAGATGGGCAAGAGTGAGA 2833
QY 2752 CTGGCCCTCTTCAAGTGTGCGAGGCTGAGTGTGAGGC 2789
Db 2834 TGCTGTGTGTGTGGCCACCGACGGGAGTGTGCTGAGCC 2871

Search completed: October 5, 2004, 15:38:48
Job time : 1125 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: October 5, 2004, 15:11:27 ; Search time 199 Seconds
(without alignments)
8034.239 Million cell updates/sec

Title: US-09-970-944-1

Perfect score: 2881

Sequence: 1 agctggggtccggctgag.....ccttccacacacgggggaga 2881

Scoring table: IDENTITY NUC

Gapop 10.0 , Gapext 1.0

Searched: 682709 seqs, 277475446 residues

Total number of hits satisfying chosen parameters: 1365418

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

Issued Patents NA: *

1: /cgn2.6/prodata/2/ina/5A_COMB.seq: *
2: /cgn2.6/prodata/2/ina/5B_COMB.seq: *
3: /cgn2.6/prodata/2/ina/6A_COMB.seq: *
4: /cgn2.6/prodata/2/ina/6B_COMB.seq: *
5: /cgn2.6/prodata/2/ina/PCTUS_COMB.seq: *
6: /cgn2.6/prodata/2/ina/backfiles1.seq: *

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	2235.8	77.6	3014	2	US-08-808-982-1
2	2235.8	77.6	3014	3	US-09-306-902A-1
3	1570.4	54.5	1787	2	US-08-808-982-2
4	1570.4	54.5	1787	3	US-09-306-902A-2
5	1147.8	39.8	1282	4	US-09-833-381-1806
6	783.4	27.2	2831	2	US-08-808-982-3
7	783.4	27.2	2831	3	US-09-306-902A-3
8	419	14.5	1605	4	US-09-833-381-1807
9	256	8.9	771	1	US-08-253-155A-17
10	119	4.1	305	2	US-08-808-982-4
11	119	4.1	305	3	US-09-306-902A-4
12	51.8	1.8	1280	3	US-08-483-533-38
13	51.8	1.8	1280	4	US-09-283-471A-38
14	51.2	1.8	699	4	US-09-252-991A-7947
15	51.2	1.8	1302	4	US-09-252-991A-7947
16	51.2	1.8	1947	4	US-09-252-991A-7533
17	50.8	1.8	1770	4	US-09-252-991A-7809
18	50.8	1.8	3885	4	US-09-369-364A-16
19	50.8	1.8	4884	4	US-09-252-991A-12126
20	50.8	1.8	4884	4	US-09-252-991A-12126
21	50.2	1.7	1300	3	US-08-483-533-39
22	50.2	1.7	1300	4	US-09-283-471A-39
23	48.6	1.7	1068	4	US-09-252-991A-9933
24	48.6	1.7	1251	4	US-09-252-991A-10142
25	48	1.7	4257	2	US-08-690-473-1
26	48	1.7	4257	3	US-09-259-821A-1
27	48	1.7	4257	3	US-08-843-659-1

c	28	48	1.7	8312	4	US-09-620-312D-1048	Sequence 1048, Ap
	29	47.8	1.7	1288	1	US-08-440-856A-9	Sequence 9, Appli
c	30	47.8	1.7	2211	4	US-09-252-991A-11256	Sequence 11256, A
	31	47.8	1.7	2289	4	US-09-252-991A-10995	Sequence 10995, A
	32	47.6	1.7	1627	4	US-09-675-018B-11	Sequence 11, Appl
	33	47	1.6	7218	1	US-08-312-463-14	Sequence 14, Appl
	34	46.8	1.6	1951	3	US-08-222-865-1	Sequence 1, Appli
	35	46.8	1.6	1951	4	US-09-510-949-1	Sequence 1, Appli
	36	46.4	1.6	453	4	US-09-252-991A-4004	Sequence 3971, Ap
	37	46.4	1.6	567	4	US-09-252-991A-4004	Sequence 4004, Ap
	38	46.2	1.6	11220	3	US-09-105-537-5	Sequence 32, Appl
	39	46.2	1.6	36778	3	US-09-105-537-5	Sequence 5, Appli
	40	46.2	1.6	38506	3	US-09-320-878-19	Sequence 19, Appl
	41	46.2	1.6	38506	4	US-09-141-908-1	Sequence 19, Appl
	42	46.2	1.6	38506	4	US-09-657-440-19	Sequence 31, Appl
	43	46	1.6	1065	4	US-09-340-798A-31	Sequence 6, Appli
	44	46	1.6	3624	1	US-07-951-715A-6	Sequence 6, Appli
	45	46	1.6	3624	2	US-08-459-448A-6	Sequence 6, Appli

ALIGNMENTS

RESULT 1

US-08-808-982-1
; Sequence 1, Application US/08080982
; Patent No. 5939271
; GENERAL INFORMATION:
; APPLICANT: Tessier-Lavigne, Marc
; APPLICANT: Leonardo, E. David
; APPLICANT: Hink, Lindsay
; APPLICANT: Masu, Masayuki
; APPLICANT: Kazuko, Keino-Masu
; TITLE OF INVENTION: Netrin Receptors
; NUMBER OF SEQUENCES: 8
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: SCIENCE & TECHNOLOGY LAW GROUP
; STREET: 268 BUSH STREET, SUITE 3200
; CITY: SAN FRANCISCO
; STATE: CALIFORNIA
; COUNTRY: USA
; ZIP: 94104
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent In Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/808,982
; FILING DATE:
; CLASSIFICATION: 530
; ATTORNEY/AGENT INFORMATION:
; NAME: OSMAN, RICHARD A
; REGISTRATION NUMBER: 36,627
; REFERENCE/DOCKET NUMBER: UC96-217
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (415) 343-4341
; TELEFAX: (415) 343-4342
; INFORMATION FOR SEQ ID NO: 1:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 3014 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: double
; TOPOLOGY: linear
; MOLECULE TYPE: cDNA
; US-08-808-982-1

Query Match 77.6%; Score 2235.8; DB 2; Length 3014;
Best Local Similarity 88.5%; Pred. No. 0;
Matches 2474; Conservative 0; Mismatches 312; Indels 11; Gaps 4;

OY 87 ATGGCCGTCGGCCCGCTGTGGCGCTGTCGTCATGCTCGCGCTGGCTC 146
|||||

2304 GGCACGACGGTACTTGTACCTGACCTTACCTCGAGCGTGTGAGCCGACCTAGT 2363
Db |||||
2215 GGCACCCAGCAGTATCTGCACTGACCTTACCTCGAGCGCATCAAGCCGACACAGC 2274
QY |||||
2364 GACCTGGCTGCAAGCTGTGGTGTGGAGTGGAGGGGACGGGAGAGCTTCAACATC 2423
Db |||||
2275 GACCTGGCTGCAAGCTGTGGTGTGGAGTGGAGGGGATGGGAGAGCTTCAACATC 2334
QY |||||
2424 AACTTCAACATCACCAAGGACACAAGTTTGTGAGCTGTGAGTGGAGTGAAGCG 2483
Db |||||
2335 AACTTCAACATCACCAAGGACACAAGTTTGTGAGCTGTGAGTGGAGTGAAGCG 2394
QY |||||
2484 GGGTCCCGACCTGTGTGGCCCGCAGTGTGCTTCAAGATCCCTTCTCATTTGGCAGAG 2543
Db |||||
2395 GGGTCCCGACCTGTGTGGCCCGCAGTGTGCTTCAAGATCCCTTCTCATTTGGCAGAG 2454
QY |||||
2544 ATAAATTCAGCTGTGACCCACCTGTAGGGGGTGGCAGTGGGGGACTTGGCCGAG 2603
Db |||||
2455 ATCATGGCCAGTGTGACCCACCTGTGAGCCGGGGCGGACTGGAGAACTCTAGCCAG 2514
QY |||||
2604 AAATCCACCTGTGACAGCCATCTCAGCTTCTTTGCTTCAAGCCGACCCACAGCCATG 2663
Db |||||
2515 AAATTCACCTGTGACAGCCATCTTGTGCTTCAAGCCGACCCACAGCCATG 2574
QY |||||
2664 ATCTCACTGTGGAGCGGGCACTTCCCAAGCGCAACTCAGCCAGCTGGCTGCA 2723
Db |||||
2575 ATCTCACTGTGGAGCGGGCACTTCCCAAGCGCAACTCAGCCAGCTGGCTGCA 2634
QY |||||
2724 GCAGTGGCTGAGTGGCCAGCAGACGTGGCTTTCACAGTGGAGGCTGAGTGC 2783
Db |||||
2635 GCTGTGGCCGAGTGGCCCAACAGATGCTGGCTTTCACAGTGGAGGCGGAGTGT 2694
QY |||||
2784 TGAGCCGCGCCAGGCGCG--ACACCTACACTCTCAGCAGCTTTGGCACCACCAAGGACA 2841
Db |||||
2695 TGAGACCGACGAGCGCGTAAATGCTTACATTTCTACAGCTTGGACCTGCCAGGACA 2754
QY |||||
2842 GGCAGAGCGGACAGGGGCGCTTCCCAACCGGGG 2878
Db |||||
2755 GGCAGAGCGGACAGGGGCGCTTCCCAACCGGGG 2791

RESULT 2

US-09-306-902A-1

; Sequence 1, Application US/09306902A

; Patent No. 6277585

; GENERAL INFORMATION:

; APPLICANT: Tessier-Lavigne, Marc

; Leonardo, E. David

; Hink, Lindsay

; Masu, Masayuki

; Kazuko, Keino-Masu

; TITLE OF INVENTION: Netrin Receptors

; NUMBER OF SEQUENCES: 9

; CORRESPONDENCE ADDRESS:

; ADDRESSEE: SCIENCE & TECHNOLOGY LAW GROUP

; STREET: 268 BUSH STREET, SUITE 3200

; CITY: SAN FRANCISCO

; STATE: CALIFORNIA

; COUNTRY: USA

; ZIP: 94104

; COMPUTER READABLE FORM:

; MEDIUM TYPE: Floppy disk

; COMPUTER: IBM PC compatible

; OPERATING SYSTEM: PC-DOS/MS-DOS

; SOFTWARE: PatentIn Release #1.0, Version #1.30

; CURRENT APPLICATION DATA:

; APPLICATION NUMBER: US/09306,902A

; FILING DATE: 07-May-1999

; CLASSIFICATION: <Unknown>

; ATTORNEY/AGENT INFORMATION:

; NAME: OSMAN, RICHARD A

; REGISTRATION NUMBER: 36,627

; REFERENCE/DOCKET NUMBER: UC96-217

TELECOMMUNICATION INFORMATION:

; TELEPHONE: (415) 343-4341

; TELEFAX: (415) 343-4342

; INFORMATION FOR SEQ ID NO: 1:

; SEQUENCE CHARACTERISTICS:

; LENGTH: 3014 base pairs

; TYPE: nucleic acid

; STRANDEDNESS: double

; TOPOLOGY: linear

; MOLECULE TYPE: cDNA

; SEQUENCE DESCRIPTION: SEQ ID NO: 1:

US-09-306-902A-1

Query Match

Best Local Similarity 77.6%; Score 2235.8; DB 3; Length 3014;

Matches 2474; Conservative 0; Mismatches 312; Indels 11; Gaps 4;

QY	87	ATGGCGCTCCGCGCGCTGTGGCAGCGCTTGGGCGATAGTCTCGCGCTTGGCTC	146
Db	1	ATGGCGCTCCGCGCGCGCTGTGGCAGCGCTTGGGCGATAGTCTCGCGCTTGGCTC	60
QY	147	CGCGCTCGGCTGCGCGCGCTGTGGCAGCGCTTGGGCGATAGTCTCGCGCTTGGCTC	206
Db	61	CGTGGTTGGGCTGCGCGCGCTGTGGCAGCGCTTGGGCGATAGTCTCGCGCTTGGCTC	120
QY	207	GACCTGCTTCCCGCTTCTGTGGAGCGCGGAGGATGTGATCATGTCAAGAACAGCCA	266
Db	121	GACCTGCTTCCCGCTTCTGTGGAGCGCGGAGGATGTGATCATGTCAAGAACAGCCA	180
QY	267	GTGCTGTGTGTGCAAGGCGCTGCGCGCGCTGTGGCAGCGCTTGGGCGATAGTCTCGCGCTTGGCTC	326
Db	181	GTGCTGTGTGTGCAAGGCGCTGCGCGCGCTGTGGCAGCGCTTGGGCGATAGTCTCGCGCTTGGCTC	240
QY	327	TGGGTCGCGCGCTGTGGCAGCGCTTGGGCGATAGTCTCGCGCGCTTGGGCGATAGTCTCGCGCTTGGCTC	386
Db	241	TGGGTCGCGCGCTGTGGCAGCGCTTGGGCGATAGTCTCGCGCGCTTGGGCGATAGTCTCGCGCTTGGCTC	300
QY	387	ACCATGGAGTCCGCAATTAATGTCTCAAGGCGAGCGAGTGTGGGCTGGAG	446
Db	301	ACCATGGAGTCCGCAATTAATGTCTCAAGGCGAGCGAGTGTGGGCTGGAG	360
QY	447	GAATACCTGTGCGCTGTGGCAGCGCTTGGGCGATAGTCTCGCGCGCTTGGGCGATAGTCTCGCGCTTGGCTC	506
Db	361	GAATACCTGTGCGCTGTGGCAGCGCTTGGGCGATAGTCTCGCGCGCTTGGGCGATAGTCTCGCGCTTGGCTC	420
QY	507	TACATCCGCAATTCGCGCGCTGTGGCAGCGCTTGGGCGATAGTCTCGCGCGCTTGGGCGATAGTCTCGCGCTTGGCTC	566
Db	421	TACATCCGCAATTCGCGCGCTGTGGCAGCGCTTGGGCGATAGTCTCGCGCGCTTGGGCGATAGTCTCGCGCTTGGCTC	480
QY	567	TCCCTGGAGCGAGGATCGTGTGGCAGCGCTTGGGCGATAGTCTCGCGCGCTTGGGCGATAGTCTCGCGCTTGGCTC	626
Db	481	TCCCTGGAGCGAGGATCGTGTGGCAGCGCTTGGGCGATAGTCTCGCGCGCTTGGGCGATAGTCTCGCGCTTGGCTC	540
QY	627	GTGAGTGGCTCCGCAATTAATGTCTCAAGGCGAGCGAGTGTGGGCTGGAG	686
Db	541	GTGAGTGGCTCCGCAATTAATGTCTCAAGGCGAGCGAGTGTGGGCTGGAG	600
QY	687	ACGCGGAGCAGCGCTGTGGTGTGGCAGCGCGCTTGGGCGATAGTCTCGCGCGCTTGGGCGATAGTCTCGCGCTTGGCTC	746
Db	601	ACGCGGAGCAGCGCTGTGGTGTGGCAGCGCGCTTGGGCGATAGTCTCGCGCGCTTGGGCGATAGTCTCGCGCTTGGCTC	660
QY	747	TGCGTGGCAGCAATCGTGTGGCAGCGCTTGGGCGATAGTCTCGCGCGCTTGGGCGATAGTCTCGCGCTTGGCTC	806
Db	661	TGCGTGGCAGCAATCGTGTGGCAGCGCTTGGGCGATAGTCTCGCGCGCTTGGGCGATAGTCTCGCGCTTGGCTC	720
QY	807	GTGAACGCTGGTGTGGTGTGGCAGCGCGCTTGGGCGATAGTCTCGCGCGCTTGGGCGATAGTCTCGCGCTTGGCTC	866
Db	721	GTGAACGCTGGTGTGGTGTGGCAGCGCGCTTGGGCGATAGTCTCGCGCGCTTGGGCGATAGTCTCGCGCTTGGCTC	780
QY	867	GGCTGGCAGCAATCGTGTGGCAGCGCGCTTGGGCGATAGTCTCGCGCGCTTGGGCGATAGTCTCGCGCTTGGCTC	926
Db	781	GGCTGGCAGCAATCGTGTGGCAGCGCGCTTGGGCGATAGTCTCGCGCGCTTGGGCGATAGTCTCGCGCTTGGCTC	840

QY 2788 GCGGCCAGGCCGACACTACACTCTCACAGCTTTGGCA--CCCAAGGACAGCA 2845
DB 1672 GCGGCCAGG--CGAACACTACATTTTACAGTTTGGGAACCCCAAGGACAGCA 1729
QY 2846 GAAGCCGACAGGGCCCTTCCACACCGGGGAGA 2881
DB 1730 GAAGCCGACAGGGCTTTTCCAAAACCGGGGAGA 1765

RESULT 4

US-09-306-902A-2

; Sequence 2, Application US/09306902A

; Patent No. 6277585

; GENERAL INFORMATION:

; APPLICANT: Tessier-Lavigne, Marc

; Leonardo, E. David

; Hink, Lindsay

; Masu, Masayuki

; Kazuko, Keino-Masu

; TITLE OF INVENTION: Netrin Receptors

; NUMBER OF SEQUENCES: 9

; CORRESPONDENCE ADDRESS:

; ADDRESS: SCIENCE & TECHNOLOGY LAW GROUP

; STREET: 268 BUSH STREET, SUITE 3200

; CITY: SAN FRANCISCO

; STATE: CALIFORNIA

; COUNTRY: USA

; ZIP: 94104

; COMPUTER READABLE FORM:

; MEDIUM TYPE: Floppy disk

; COMPUTER: IBM PC compatible

; OPERATING SYSTEM: PC-DOS/MS-DOS

; SOFTWARE: Patent In Release #1.0, Version #1.30

; CURRENT APPLICATION DATA:

; APPLICATION NUMBER: US/09/306,902A

; FILING DATE: 07-May-1999

; CLASSIFICATION: <Unknown>

; ATTORNEY/AGENT INFORMATION:

; NAME: OSMAN, RICHARD A

; REGISTRATION NUMBER: 36,627

; REFERENCE/DOCKET NUMBER: UC96-217

; TELECOMMUNICATION INFORMATION:

; TELEPHONE: (415) 343-4341

; TELEFAX: (415) 343-4342

; INFORMATION FOR SEQ ID NO: 2:

; SEQUENCE CHARACTERISTICS:

; LENGTH: 1787 base pairs

; TYPE: nucleic acid

; STRANDEDNESS: double

; TOPOLOGY: linear

; MOLECULE TYPE: cDNA

; SEQUENCE DESCRIPTION: SEQ ID NO: 2:

US-09-306-902A-2

Query Match

Best Local Similarity 54.5%; Score 1570.4; DB 3; Length 1787;

Matches 1721; Conservative 0; Mismatches 36; Indels 19; Gaps 11;

QY 1114 GCAACTGTACCACTCTGTGTACACAGTCTTCTGGCCCTTGAGGAGTGGCCCTCT 1173
DB 1 GCNACTGTACCACTGTACCTCTG-GTACACTCTTCTGGCCCTTGAGGAGTGGCCCTCT 59
QY 1174 ATGTGGCCCTATCGCGGTGGCCGTCTGCTGTGCTGTGCTGTGCTGTGCTATCTCTG 1233
DB 60 ATGTGGCCCTATCGCGGTGGCCGTCTGCTGTGCTGTGCTGTGCTGTGCTATCTCTG 119
QY 1234 TTTATTGCGGAAGAGAGGGGCTGGACTCAGATGGGCTGACTCGTCCATTCTCACT 1293
DB 120 TTTATTGCGGAAGAGAGGGGCTGGACTCAGATGGGCTGACTCGTCCATTCTCACT 179
QY 1294 CAGGCTTCCAGCCCGTTCAGATC-AAGCCGAGCAAGCAGCAACCCCAATCTGCTCACC 1352

DB 180 CAGGCTTCCAGCCCGTTCAGCATCTAAGCCCAAGCAGCAACCCCACTCTGCTCACC 239
QY 1353 ATCCAGCCGGACCTCAG--CACCACACCACTTACAGGGCAGTCTCTGTCCCGGAG 1409
DB 240 ATCCAGCCGGACCTCAGCACCACCACTTACAGGGCAGTCTCTGTCCCGGAG 299
QY 1410 GATGGGCCAGCCCAAGTTCCAGCTCACCATGGGCACCTGCTCAGCCCCCTGGGTGGC 1469
DB 300 GATGGGCCAGCCCAAGTTCCAGCTCACCATGGGCACCTGCTCAGCCCCCTGGGTGGC 359
QY 1470 GCGGCCACACATGACACACAGCTCTCCACCTCTGAGGCCGAGGAGTGTCTCCCGC 1529
DB 360 GCGGCCACACATGACACACAGCTCTCCACCTCTGAGGCCGAGGAGTGTCTCCCGC 419
QY 1530 CTCTCCACCCAGAACTACTTCGGCTCCCTGCCCGCAGCACCAGCAACATGACCTATGG 1589
DB 420 CTCTCCACCCAGAACTACTTCGGCTCCCTGCCCGCAGCACCAGCAACATGACCTATGG 479
QY 1590 ACCTTCAACTTCTCTCGGGGCGGCTGATGATCCCTAATACAGGTATCAGCCTCTCATC 1649
DB 480 ACCTTCAACTTCTCTCGGGGCGGCTGATGATCCCTAATACAGGAATCAGCCTCTCATC 539
QY 1650 CCCCAGATGCCATACCCCGAGGAGATCTATGAGATCTACTCAGCTGCACAAGCG 1709
DB 540 CCCCAGATGCCATACCCCGAGGAGATCTATGAGATCTACTCAGCTGCACAAGCG 599
QY 1710 GAAGACGTGAGGTTCGCCCTAGCTGCTGACAGCCTCTGAGTCCCATCGTTAGCTGT 1769
DB 600 GAAGACGTGAGGTTCGCCCTAGCTGCTGACAGCCTCTGAGTCCCATCGTTAGCTGT 659
QY 1770 GGACCCCTTGGCGTCTGTCTACCCGCGCAGTCTCTGGCTATGACCACTGTGGGAG 1829
DB 660 GGACCCCTTGGCGTCTGTCTACCCGCGCAGTCTCTGGCTATGACCACTGTGGGAG 719
QY 1830 CCCCAGCTGACAGCTGGAGCCTGGCCCTCAAAAAGCAGTCTGCGAGGCGAGCTGGGAG 1889
DB 720 CCCCAGCTGACAGCTGGAGCCTGGCCCTCAAAAAGCAGTCTGCGAGGAG-AGCTGGGAG 778
QY 1890 CAGGATGTCTGACCTGGCGGAGGAGCGCCCTCCACCTCTACTACTGCCAGCTGGAG 1949
DB 779 GATG---TCTGACCTGGCGGAGGAGCGCCCTCCACCTCTACTACTGCCAGCTGGAG 834
QY 1950 GCGAGTGTCTGCTTCTTCCAGGAGCTGGGCGCTTTGCGCTGTGGGAGAGGCC 2009
DB 835 GCGAGTGTCTGCTTCTTCCAGGAGCTGGGCGCTTTGCGCTGTGGGAGAGGCC 894
QY 2010 CTCAGCTGTGCTGCGCCAAAGCGCTCAAGCTGCTTCTGTTGCGCGGCTGGCCTGCACC 2069
DB 895 CTCAGCTGTGCTGCGCCAAAGCGCTCAAGCTGCTTCTGTTGCGCGGCTGGCCTGCACC 954
QY 2070 TCCCTCGAGTACAAATCCGGGTCTACTGCTGCATGACACCCAGATGCACCTCAAGGAG 2129
DB 955 TCCCTCGAGTACAAATCCGGGTCTACTGCTGCATGACACCCAGATGCACCTCAAGGAG 1014
QY 2130 GTGGTGCAGCTGGAGAAGCAGCTGGGGGACAGCTGATCCAGAGCAGCGGTCTGAC 2189
DB 1015 GTGGTGCAGCTGGAGAAGCAGCTGGGGGACAGCTGATCCAGAGCAGCGGTCTGAC 1074
QY 2190 TTCAGAGAGATTTACCAACCTGCGCCTATCCATCCAGATGTGCCAGCTCCCTGTGG 2249
DB 1075 TT-AAGGACAGTTTACCACACCT--GCCCTATCATCCAGATGTGCCAGCTCCCTGTGG 1131
QY 2250 AAGAGTAAGCTCTCTGTGCTACAGTACACAGGAGATCCCTTTTATCACAATGGAATGGCAG 2309
DB 1132 AAGAGTAAGCTCTCTGTGCTACAGTACACAGGAGATCCCTTTTATCACAATGGAATGGCAG 1191
QY 2310 CAGCGTATCTGACCTTTCACCTTGAGGAGTGTGAGCCCGAGCAGTGTGAGCTG 2369
DB 1192 CAGCGTATCTGACCTTTCACCTTGAGGAGTGTGAGCCCGAGCAGTGTGAGCTG 1251
QY 2370 GCCTGCAAGCTGTGGGTGTGGCAGGTGGAGGCGAGCGGGCAGAGCTTCAGCATCACTTC 2429
DB 1252 GCCTGCAAGCTGTGGGTGTGGCAGGTGGAGGCGAGCGGGCAGAGCTTCAGCATCACTTC 1311

```
QY 2430 AACATCACCAAGGACACAAAGGTTTGTGAGTCTGTGGCTCTGGAGAGTGAAGCGGGGTC 2489
Db 1312 AACATCACCAAGGACACAAAGTGTGTGAGTCTGTGGCTCTGGAGAGTGAAGCGGGGTC 1371
QY 2490 CGAGCCCTGTGGGCCCCAGTGCCTTCAAGATCCCCCTTCTCAATCGGCAGAGATAAAT 2549
Db 1372 CCAGCCCTGTGGGCCCCAGTGCCTTCAAGATCCCCCTTCTCAATCGGCAGAGATAAAT 1431
QY 2550 TCAGAGCTGAGACCAACCCCTGTAGGCGGGGTGCCGACTGGCGGACTCTGGCCGAGAACTC 2609
Db 1432 TCAGAGCTGAGACCAACCCCTGTAGGCGGGGTGCCGACTGGCGGACTCTGGCCGAGAACTC 1491
QY 2610 CACCTGGACAGCCATCTCAGCTTCTTTTGGCTTCAAGCCAGCCAGCCACAGCATGATCCCTC 2669
Db 1492 CACCTGGACAGCCATCTCAGCTTCTTTTGGCTTCAAGCCAGCCAGCCACAGCATGATCCCTC 1551
QY 2670 AACCTGTGGAGGCGGGGCACTTCCCAACGGCAACCTCAGCCAGCTGGCTGAGCAGTG 2729
Db 1552 AACCTGTGGAGGCGGGGCACTTCCCAACGGCAACCTCAGCCAGCTGGCTGAGCAGTG 1611
QY 2730 GTTGGACTGGGCGAGCAGCAGCTGGCCTC-TTCACAGTG-TGGAGGCTGAGTGTGAG 2787
Db 1612 GTTGGACTGGGCGAGCAGCAGCTGGCCTC-TTCACAGTG-TGGAGGCTGAGTGTGAG 1671
QY 2788 GCCGGCCAGGCGGCGCACTACACTCTCAGCAGCTTGGCA--CCCAACAGGACAGGCA 2845
Db 1672 GCCGGCCAGG--CGAACACTACAAATTTTACAGTTTGGAGCCCAAGGACAGGCA 1729
QY 2846 GAAAGCGGACAGGGGCGCTTCCCAACCGGGGAGA 2881
Db 1730 GAAAGCGGACAGGGGCTTTTCCCAAAACCGGGGAGA 1765

RESULT 5
US-09-833-381-1806
; Sequence 1806, Application US/09833381
; Patent No. 6672186
; GENERAL INFORMATION:
; APPLICANT: Robison, Keith E.
; TITLE OF INVENTION: No. 6672186el Nucleic Acid and Protein Homologs
; FILE REFERENCE: 5800-119
; CURRENT APPLICATION NUMBER: US/09/833,381
; CURRENT FILING DATE: 2001-04-11
; PRIOR APPLICATION NUMBER: 09/516,448
; PRIOR FILING DATE: 2000-02-29
; NUMBER OF SEQ ID NOS: 2050
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 1806
; LENGTH: 1282
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (1)...(1282)
; OTHER INFORMATION: n = A,T,C or G
US-09-833-381-1806

Query Match 39.8%; Score 1147.8; DB 4; Length 1282;
Best Local Similarity 97.7%; Pred. No. 3.3e-240;
Matches 1259; Conservative 0; Mismatches 18; Indels 11; Gaps 9;

QY 1510 CCGAGGAGTTCGTCTCCCGCTCTCCACCCAGAACTACTCCGTCCTCGGCGCGGCTGATGATCCCTAATA 1629
Db 1 CCGAGGAGTTCGTCTCCCGCTCTCCACCCAGAACTACTCCGTCCTCGGCGCGGCTGATGATCCCTAATA 60
QY 1570 CCAGCAACATACCTATGAGGACCTTCAACTTCTCGGCGCGGCTGATGATCCCTAATA 1629
Db 61 CCAGCAACATGACCTATGAGGACCTTCAACTTCTCGGCGCGGCTGATGATCCCTAATA 120
QY 1630 CAGGTATCAGCTCTCTATCCCCCAGATGCCATACCCGAGGAGAGATCTATGAGATCT 1689
Db 121 CAGGAATCAGCTCTCTATCCCCCAGATGCCATACCCGAGGAGAGATCTATGAGATCT 180
```

```
QY 1690 ACCTCAGCTGCACAAGCCGGAAGACGTGAGGTTCGCCCTAGCTGGCTGTTCAGACCCCTGC 1749
Db 181 ACCTCAGCTGCACAAGCCGGAAGACGTGAGGTTCGCCCTAGCTGGCTGTTCAGACCCCTGC 240
QY 1750 TGAGTCCCCTATCGTTAGCTGTGACCCCTTGGCGTCTGCTTGGCGTCTGCTTGGCGTCTGCT- 1807
Db 241 TGAGTCCCCTATCGTTAGCTGTGACCCCTTGGCGTCTGCTTGGCGTCTGCTTGGCGTCTGCT- 300
QY 1808 GGCTATGAGCCACTGTGGGAGAGCCAGCCCTGACAGCT--GGAGCTTGGCGTCAAAAAGC 1866
Db 301 GGCTATGAGCCACTGTGGGAGAGCCAGCCCTGACAGCTGGAGCTTGGCGTCAAAAAGC 360
QY 1867 AGTCGTGCAGGCGAGCTGGAGAGAGTGTGCTGACCTGGGCGAGGAGGCGCCCTCC 1926
Db 361 AGTCGTGCAGGCGAGCTGGAGAGAGTGTGCT--TGCACCTGGGCGAGGAGGCGCCCTCC 418
QY 1927 ACCTCTACTACTGCGAGCTGGAGGCGAGTGCCTGCTACGTCTTTCACCGAGCAGCTGGGC 1986
Db 419 ACCTCTACTACTGCGAGCTGGAGGCGAGTGCCTGCTACGTCTTTCACCGAGCAGCTGAGCC 478
QY 1987 GCTTTGCCCTGTGGAGAGGCGCTCAGCGTGGCTGCCGCCAAGCGCCTCAAGCTGCTTC 2046
Db 479 GCTATGCCCTGTGGAGAGGCGCTCAGCGTGGCTGCCGCCAAGCGCCTCAAGCTGCTTC 538
QY 2047 TGTTCGCGCGGTGGCTGACCTCCCTCGAGTACAACATCCGGGTCTACTGCTGCATG 2106
Db 539 TGTTCGCGCGGTGGCTGACCTCCCTCGAGTACAACATCTGCTTACTGCTGCATG 598
QY 2107 ACACCCAGATGCACCTCAGGAGGTGTGACAGCTGGAGAGCAGCTGGGGGAGACAGTGA 2166
Db 599 ACACCTCAGATGCACCTCAGGAGGTGTGACAGCTGGAGAGCAGCTGGGGGAGACAGTGA 658
QY 2167 TCCAGGAGCCAGGTCCTTGCACCTTCAAGGACAGTTACCAACCTTGCCTTATCCATCC 2226
Db 659 TCCAGGAGCCAGGTCCTTGCACCTTCAAGGACAGTTACCAACCTTGCCTTATCCATCC 718
QY 2227 AGGATGCGCCAGCTCCCTGTGGAAGAGTAAGTCTTGTGACAGTACCAGGAGATCCCT 2286
Db 719 AGGATGTCCTCAGCTCCCTGTGGAAGAGTAAGTCTTGTGACAGTACCAGGAGATCCCT 778
QY 2287 TTTATCACATCTGGAATGGCAGCGGTACTTGCACCTGACCTTCACTTGGAGCGTG 2346
Db 779 TTTATCACATCTGGAATGGCAGCGGTACTTGCACCTTCACTTGGAGCGTG 838
QY 2347 TCAGCCCGCAGCACTAGTGCACCTGGCTTCAAGCTGTGGTGTGGAGGCGGAGC 2406
Db 839 TCAGCCCGCAGCACTAGTGCACCTGGCTTCAAGCTGTGGTGTGGAGGCGGAGC 898
QY 2407 GCGAGAGCTTCAGCATCAACTTCAACATCAAGGACACAAAGGTTTGTGAGCTGCTGG 2466
Db 899 GCGAGAGCTTCAGCATCAACTTCAACATCAAGGACACAAAGGTTTGTGAGCTGCTGG 958
QY 2467 CTCTGGAGAGTGAAGCGGGGTCCAGCCCTGGTGGGCGCCAGTGCCTTCAAGATCCCT 2526
Db 959 CTCTGGAGAGTGAAGCGGGGTCCAGCCCTGGTGGGCGCCAGTGCCTTCAAGATCCCT 1018
QY 2527 TCCTCATTTGGCAGAGATAA-TTTTCCAGCTTGCACCCACCTGTAGGCGGGGTGCCGAC 2585
Db 1019 TCCTCATTTGGCAGAGATAA-TTTTCCAGCTTGCACCCACCTGTAGGCGGGGTGCCGAC 1078
QY 2586 TGGCGG-ACCTTGGCGCAGAAATCTCCACCTGGACAGCACTCTAGCTTCTTTGGCTCCAA 2644
Db 1079 TGGCGGAACTCTGGCGCAGAAATCTCCACCTGGACAGCACTCTAGCTTCTTTGGCTCCAA 1138
QY 2645 GCCCAGCGCCACAGCCATGATCCCTCAACCTTGGAGGCGGCGGACATTTCCCAAGCGGAA 2704
Db 1139 GCCCAGCGCCACAGCCATGATCCCTCAACCTTGGAGGCGGCGGACATTTCCCAAGCGGAA 1198
QY 2705 CCTCAGCGAGCTGCTGACAGTGGCTGGACTGGGCGAGCAGCAGCTGCTCTTAC 2764
Db 1199 CCTCAGCGAGCTGCTGACAGTGGCTGGACTGGGCG--AGCAGACGCTG-CTCTTAC 1254
```

Qy 2765 AGTGCGGAGGCTGAGTGTCTGAGGCCGG 2792
|||
Db 1255 AGTGCGGAGGCTGAGTGTCTGAGGCCGG 1282

RESULT 6

```

US-08-808-982-3
; Sequence 3, Application US/08808982
; Patent No. 5939271
; GENERAL INFORMATION:
; APPLICANT: Tessier-Lavigne, Marc
; APPLICANT: Leonardo, E. David
; APPLICANT: Hink, Lindsay
; APPLICANT: Masu, Masayuki
; APPLICANT: Kazuko, Keino-Masu
; TITLE OF INVENTION: Netrin Receptors
; NUMBER OF SEQUENCES: 8
; CORRESPONDENCE ADDRESS:
; ADDRESSSEE: SCIENCE & TECHNOLOGY LAW GROUP
; STREET: 268 BUSH STREET, SUITE 3200
; CITY: SAN FRANCISCO
; STATE: CALIFORNIA
; COUNTRY: USA
; ZIP: 94104
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version 1
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/808.982

```

Query Match	27.2%	Score 783.4;	DB 2;	Length 2831;
Best Local Similarity	59.4%	Pred. No. 5.7e-161;		
Matches 1625; Conservative	0;	Mismatches 971;	Indels 139;	Gaps 12;

QY	184	ACCGAGTCCTGGTGCCAAACCGGACCTGCTTCCCACTTCTCGTGTGAGCCCGGAGGATG	243
Db	104	ACTCCTTCCCATCAGACCCCGGAGCAGCTGCCCTCACTTCTCGTGTGAACAGAGGATG	163
QY	244	TGTACATCGTCAAGAAACAAGCCAGTGTCTTGTGTGCAAGGCCGTGCGCCGACCGCAGA	303
Db	164	CTACATCGTAAAGAAACAAGCCAGTGTGGAATTCGACTGCCAGCCTTCCCTGCCACACAGA	223
QY	304	TCTTCTTCAAGTGCACCGGGAGTGGGTGCGCCAGGTGGAACAAGTGATTCGAGCCGACAGA	363
Db	224	TCTACTTCAAGTGTAAATGGCGAGTGGGTAGCCAGAAAGGCCACGCTCACGCGAGGAGGCC	283
QY	364	CAGACGGGAGCAGTGGTGAGCCGACCATCGGAGGTCCGCATTAATGTCTCAAGCGCAGCAGG	423
Db	284	TGGATGAGCCACAGCTTTCGAAATCAGAGAGTGCAGATAGAGTGTGCGCGCAGCAGG	343
QY	424	TGCAGAAGGTGTTCGGGCTGGAGGAATACTGTGTCCAGTGCCTGGGATAGGAGCTCTCTCGG	483
Db	344	TGGAGAACTTTTGTGGCTTCGAGGACTACTGTGTTCAGTGCCTGGGCTGGAGCTCTTCGG	403

164 CCTCATCGTAAAGAACAGCCAGTGGAAATTCACATGCGGAGCCTTCCCTGCCACACAGA 223
1304 TCTCTTTCAAGTGAACAGGGAGTGGTGGCCAGGTGGACACAGTGTGAGCGCAGCA 363
224 TCTACTTCAAGTGTAAATGGCAGTGGTTAGCCAGAAAGGCCACGTACAGCAGGAGAGCC 283
364 CAGCGGAGAGTGGTGAAGCCAGCAGATGGAGGTTCGCATTAATGTCTCAAGGAGCAGG 423
284 TGGATGAGGCCACAGGCTTGGCAATACGAGAGGTGCAGATAGAGGTGTGCGGCGCAGG 343
424 TCGAGAAGTGTTCGGCTGGAGGAATACTGGTGCAGTGGCTGGCATGGAGCTCCTCGG 483
344 TCGAGAAGTGTTCGGCTGGAGGACTACTGGTGTGAGTGGCTGGAGCTCTTCGG 403
484 GCACCAACAGAGTCAGAAAGGCTACATCCGATAGCCAGATGGCGAAGAACTTCGAGC 543
404 GAAACCAAGAGTCGCGAGGCTACATCCGATTCGCTTACTTGGCAAGAACTTTGACC 463
544 AGGAGCGCTGCCAAGGAGTGTCCCTGGAGCAGGGCATCGTGTGCCCTCGGCTCCAC 603
464 AGGAGCCTTGGCGAAGGAGGTACCTTGGATCATGAGGTCTTCTGCGAGTGGCGCCAC 523
604 CGGAGGCACTCCTCAGCGGAGGTGGAGTGGCTCGGAACGAGGACCTGGTGGACCCGT 663
524 CAGAGGAGTGTCTGGCTGAGTGGATGGCTCAAGAAATGAAGATGTATCGATCCCG 583
664 CCTCGACCCCAATPATATCATCAGCGGGAGCAGAGCTGGTGGTGGACAGGCGCGCC 723
584 CTCAGGACACTAACTTCCTGCTCACCATTGACCAACCTCATCTCCGCGAGGCGCGCC 643
724 TTGCTGACACGGCAACTACCTGCTGGTGGCCAGAACATCGTGACCTGCGCGCAGCG 783
644 TCTCAGACACGCCAACTACCTGTGTGGCAAGAAATATTTGGTCAAGCGCGCGGAGCA 703
784 CTTCCGCTCTGTCTATCGTCTAGTGAACGGTGGTGGTTCGAGTGGACGAGTGGTCCG 843
704 CGACGGCAAGTATCTGTATGTGAACGGAGTTGGTCCAGTGGCGAGATGTGTAC 763
844 TCTCAGCGCCCACTGTGGCGCGCTGGCAGAAACGAGCGGAGCTGACCAACCGCGG 903
764 CTTGCTTAACCGCTCGCGCGAGGTGGCAGAAAGTATAGGACCTGACCAACCGCAG 823
904 CGCTCTCAACGGGGCGCTTCTGTGAGGGGAGAAATGTCAATGACCGCAGCCTCTCT 963
824 CCCCCTCAATGAGGTGGCTTCTGCGAGGAGCAGGCTTGCCAGAA---GACGGCTTGA 880
964 CTCTGCTGTCTGTGTGAGCGGAGCTGGAGCCCGTGGAGCAAGTGGTGGCCCTGTGGCC 1023
881 CCACCGTGTGCCAGTGGATGGAGCGTGGACTGAGTGGAGCAAGTGGTCCGCTGCAGCA 940
1024 TGGACTGCACCCACTGGCGAGCGGTGAGTGTCTGACCCAGACCGCGCGACCGAGGGG 1083
941 CAGAGTGTGGCACTGGCGAGCGCGAGTGCATGGCAGCCGCGCGCCAGCAAGCGAGGCC 1000
1084 AGGAGTGCAGGCACTGACCTGACACCGCAACTGTATACAGTACCTCTGTGTACACA 1143
1001 GTGACTGCGAGGAGCGTACTTGTACTCCAGAACTGCACCGATGGCTGTGGTGTGCTGA 1060
1144 GTGCTTTCTGGCC-----CTGAGGAGCTGGCCC 1170
1061 ATCAGAGAACTTAAGACCTTAAGCGCGCCCTGGAGCGCTGGGAGAGCTGGCGC 1120
1171 TCTATGAGCGCTC---ATGCGCGTGGCGGTCTGCGCTGGTCTGCTGCTCTCTCA 1227
1121 TGTATGCGGCGCTCGTGGTGGCGCTCTTTGTGGTCTGGCAGTTCATGGCTGTAGGAG 1180
1228 TCTCTGTTTATTCGCGAAGAGAGGGGCTGACTCAGATGTGGTGTGACTCGTCCATT- 1286
1181 TGAATGTTACCGGAGAACTGCGGGGACTTCGACACGGACATCACTGACTCTCTCTGCTG 1240
1287 --CTCAGCTCAGGCTTCAGCCCGTTCAGCATCAAGCCCGCAGCAAGCAGCAACCCCGATC 1344

1241 CCTCACTGGTGGTTTCCACCCCGCTCAAACTTCAAGACTGCAAGGCCAGCAACCCACAGC 1300
1345 TGCT-----CACCATCCAGCGGAGCCTCAGACCA---CCACCACTTACCAGGCGAGTC 1395
1301 TCTTCGACCCCATCCGCCCTCCGGACCTAAACGCGAGTGTGGCATCTACCGCGGACCTG 1360
1396 TCTGTCCCGCGGAGCA-----TACTTCCGGT 1411
1361 TGTATGCTCTGAGAGCTCTGCCGACAAGATCCCTATGACTAATTCACCCCTTCTGGATC 1420
1412 -TGGGCCCAGGCCCAAGTTCAGCTCAACCAATGGGCACTGTCTCAGCCC-----CC 1461
1421 CCTTCCCGAGCCTCAAGATCAAGTCTATGACTCCAGCACCATCGCTCTGGGGTGGCC 1480
1462 TGGGTGGGGCGCGCACACACTGCACACAGCTCTCCACCTCTCAGGGCGGAGGAGTTCG 1521
1481 TGGCTGATGGAGCCAGCTGTGGGTGTCTTACCAACCGGTATACCCAGGCGATTTCT 1540
1522 TCTCCGCGCTCTCCACCCAGAAC-----TACTTCCGGT 1554
1541 CCGGGACACCCACTTCTGTGCACCTGGCGAGCGCAGCCTTGGTTCCCGACCTCTCTGG 1600
1555 CCTTCCCGGAGGCAACAGCAATGAACTATGGGACCTTCAACTTCTTGGGGGCGCGC 1614
1601 GCTTCCCTCGAGACCCCGAGCAGTGTCAAGTGGCACCTTTGGTTGCTGGGTGGAGGC 1660
1615 TGATGATCCCTAATACAGGTATCAGCTCTCATCCCCCAGATGCATACCCCGAGGGA 1674
1661 TGACCAATTCGCGGACAGGGTCAAGCTTGTGGTACCAATGGAGCCATTCGCCAGGGA 1720
1675 AGATCTATGAGATCTACCTCAGCTGCACAAGCCGGAAGACGTGAGGTTGCCCTAGCTG 1734
1721 AGTCTATGACTGTATCTACGTATCAACAAGACTGAAAGCACCTTCCCACTTTCGAAG 1780
1735 GCTGTGACACCTGTGAGTCCCATCTAGCTGTGGACCCCTGCGCTGCTCTCACCC 1794
1781 GTTCCAGACAGTATTTAGGCCCTCGGTGACCTGCGGGGCCAGGGGCTCTCTCTGTC 1840
1795 GCGCAGTCACTCTGGTATGAGCACTGTGGGAGCCAGCCTGACGTGAGGCTGTC 1854
1841 GCGCTGTGTCTCTACTGTGCGCCCTGCTGTGAAGTCAATGCCGAGAGTGGATCTTC 1900
1855 GCTCAAAAAGCAGTCTGCGAGGCGAGTGGGAGCAGGATGTGTCACCTTGGGCGAGG 1914
1901 AGCTCAAGACCCAGGCCCATCAGGCCACTGGG---AGGAGTGTGTGACTTTGGATGAGG 1957
1915 AGCGCCCTCCACCTCTACTACTGACGTGAGGCGAGTGCCTGCTAGCTTCTCACCG 1974
1958 AGACTCTGAACACCCCTCTACTGCGAGCTAGAGCTAAATCCTGCCACATCCTGTGG 2017
1975 AGCAGCTGGCGCGCTTTGCCCTGGTGGGAGAGCCCTCAGCGTGGCTGCGCGCAAGCGCC 2034
2018 ACCAGTGGGTACTAGTGTTCACGGGCGAGTCTCTACTCCCGCTCGCAGTCAAGGGC 2077
2035 TCAAGTGTCTTGTGTGGCGCGGTGGCTGCACTCTCCCTCGAGTACAACATCCGGGTCT 2094
2078 TCCAGCTAGCACTTCTGCGCCCGAGCCCTCTGCACTCTCTGGAGTATAGTCTCAGGGTCT 2137
2095 ACTGCTGTATGACACCCAGGATGCACTCAAGAGGTTGGTGGAGTGGAGGAGCAGCTGG 2154
2138 ACTGTCTGAGGACACTCTCTGACAGCTAAGAGGTCTTAGAGCTGGAGAGCTCTGG 2197
2155 GGGGACAGCTGATCCAGGAGCCACGGGTCTGCACTTCAAGGACAGTATACCAACCTGC 2214
2198 GTGGCTACTTGTGGAGGAGCCCAAGACTTTGCTCTTTAAGGACAGTATACCAACCTAC 2257
2215 GCTATCCATCCACAGATGGCCAGTCTCCCTGTGGAGAGTAAAGTCTTGTAGCTTACC 2274
2258 G-CTCTCCCTCCATGACATCCCGCATGCCACTGGAGGAGCAAACTACTGGCCAAGTACC 2316
2275 AGGAGATCCCTTTTATCATCTGGAATGGACGCGAGCGGTACTTCTGACTGCACTTCA 2334
2317 AGGAGATTCCTTCTTACCATGTGTGGAACCGGAGCCGAGAAAGCCCTGCACTGCACTTCA 2376


```
QY 2335 CCTGGAGGCTGTACGCCCGACACCTAGTGAACCTGGCCCTGCAAGCTGTGGTGTGGCAGG 2394
|||
Db 2377 CCTGGAGAGACATAGCCTAGCCTCCTAGCTTACCTGTGAAGTCTGTGGTGGCAGG 2436
|||
QY 2395 TGGAGGGGACGGGAGAGGTTTACAGATCAATCAATTTCAACATCAC--CAAGGACACAAGGT 2451
|||
Db 2437 TAGAAGGGGAAGCCAGATTTTCCAGCTGCACACACACGCTGGGTGAGACGCTGTGGCT 2496
|||
QY 2452 TTGCTCAGCTGTGGCTCTGGAGAGTGAAGCGGGGTCCAGCCCTGGTGGGCCCCAGTG 2511
|||
Db 2497 CCTGGATGCACCTCTGCTGTGCCCTTGGCAATGCTGCCACACACAGCTGGGACCCATG 2556
|||
QY 2512 CCTCAAGATCCCTTCTCATTTGGCGAGAGATATTTCCAGCCCTGGACCCACCCCTGTA 2571
|||
Db 2557 CCTCAAGATACCAGCTGTCCATCCGCCAGAGATCTGCAACAGCTGGAGGCCCCCACT 2616
|||
QY 2572 GGGGGGTGCGACCTGCGGAGCTCTGGCCGACAAATCCACCTGGACAGCCATCTCAGCT 2631
|||
Db 2617 CACGGGGCAATGACTGGCGGTGTTGGCACAGAAAGCTCTCCATGGACCGGTACCTGAACT 2676
|||
QY 2632 TCTTTGCCCTCCAAAGCCAGCCACAGCCATGATCTCAACCTGTGGGAGGCGCGCACT 2691
|||
Db 2677 ACTTGGCCACCAAGCTAGTCCACAGGCGTGTCTTAGACCTCTGGGAAGCTCGGCAGC 2736
|||
QY 2692 TCCCAACGGCAACCTCAGCCAGCTGGCTGCAAGCAGTGGCTGACTGGGCCAGCCAGACG 2751
|||
Db 2737 AGGATGATGGGACCTCAACAGCTGGCCAGTGCCTTGGAGAGATGGCAAGAGTGAGA 2796
|||
QY 2752 CTGGGCTCTTACAGTGTGGAGGCTGAGTGCTGA 2786
|||
Db 2797 TGCTGTAGCCATGACCACTGATGGCGATTGCTGA 2831
|||
```

RESULT 8

```
US-09-833-381-1807
; Sequence 1807, Application US/09833381
; Patent No. 6672186
; GENERAL INFORMATION:
; APPLICANT: Robison, Keith E.
; TITLE OF INVENTION: No. 6672186el Nucleic Acid and Protein Homologs
; FILE REFERENCE: 5800-119
; CURRENT APPLICATION NUMBER: US/09/833,381
; CURRENT FILING DATE: 2001-04-11
; PRIOR APPLICATION NUMBER: 09/516,448
; PRIOR FILING DATE: 2000-02-29
; NUMBER OF SEQ ID NOS: 2050
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 1807
; LENGTH: 1605
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (1)...(1605)
; OTHER INFORMATION: n = A,T,C or G
US-09-833-381-1807
```

```
Query Match 14.5%; Score 419; DB 4; Length 1605;
Best Local Similarity 62.4%; Pred. No. 6.6e-82;
Matches 737; Conservative 0; Mismatches 435; Indels 10; Gaps 5;
```

```
QY 1556 CTGCCCCGAGGACACAGACATGACCTATGGGACCTTCAACTTCTCGGGGCGGCT 1615
|||
Db 421 CTGCCCCGAGACCCAGGAGCAGCGTCAGCGGACCTTTGGCTCCCTGGGTGGAGGCT 480
|||
QY 1616 GATGATCCCTAATACAGTATCAGCTCTCTATCCCCCAGATGCCATACCCCGAGGAA 1675
|||
Db 481 CANCATCCCGGACACAGGGTCAGCTTGTGTGTCCTAATGAGGCCATTCCTCCAGGCAA 540
|||
QY 1676 GATCTATCA--GATCTACTCAGCTGCACAAAGCCGGAAGAGCTGAGTTGCCCTAGCT 1733
|||
Db 541 GTTCTACGAAGATGTATTCTACTATCAACAGGCGAGAAAGTACCCTCCCGCTTTCAGAA 600
|||
```

RESULT 9
US-08-253-155A-17/c
; Sequence 17, Application US/08253155A

```
QY 1734 GGCTGTACACCCCTGCTGATGCCATCGTGTAGTGTGGACCCCTGGCGTCTGTCTCACC 1793
|||
Db 601 GGGACCAACACAGTATTAGACCCCTCGGTGACCTGTGGACCCACAGCCCTCTGTCTGTG 660
|||
QY 1794 CGGCCAGTCACTCTGGCTATGACCACTGTGGGAGCCAGCCCTGACAGCTGGAGCCCTG 1853
|||
Db 661 CGCCCCGTCACTCTCACTATGCCCCACTGTGCCGAAGTCAGTGCCCTGACTGATCTTT 720
|||
QY 1854 CGCCTCAAAAAGCAGTCTGTGGAGGSCAGCTGGGAGCAGGATGTGTCACCTTGGCGAG 1913
|||
Db 721 CAGCTCAAGACCCAGGCCCACTGGG--ANGAGTGTGTGACCCCTGGATGAG 777
|||
QY 1914 GAGGCGCCCTCCCACTCTACTACTGTCAGCTGGAGGCCAGTCCCTGCTAGCTTTCACC 1973
|||
Db 778 GAGACCTTGAACACACCCCTGCTACTGTCAGCTGGAGCCAGGGCCTGTACATCTCTGTG 837
|||
QY 1974 GAGCAGCTGGGCGGCT--TTGCCCTGTGGGAGAGGCCCTCAGCTGGCTGCCGCAAGCG 2032
|||
Db 838 GACCAGCTGGGCACCTACCGTGTTCAGGGCAGTCTTATTCGCGCTCAGCAGTCAAGCG 897
|||
QY 2033 CTTCAAGCTGTCTTGTGTGGCCCGTGGCTTGCACCTCTCGAGTACAAATCGGGT 2092
|||
Db 898 GCTCCAGCTGGCGGTCTTGGCCCCGCTCTGCACTCCCTGGAGTACAGCTCCGGT 957
|||
QY 2093 CTACTGCTTCATGACACCCAGATGCACTCAAGAGGTGTGTCAGCTGGAGAGAGCT 2152
|||
Db 958 CTACTGCTTCGAGGACACGCTGTAGCACTGAAGAGGTGTGTCAGCTGGAGAGAGCT 1017
|||
QY 2153 GGGGGACAGCTGATCCAGGAGCCACGGTCTCGTCACTTCAAGGACAGTTACCAACCT 2212
|||
Db 1018 GGGCGGATACTTGTGGAGGAGCCGAAACCGCTAATGTTCAAGGACAGTTACCAACCT 1077
|||
QY 2213 GGGCTATCCATCCACAGATGTGCCAGCTCCCTGTGGAAGAGTAAAGTCTTGTGAGTA 2272
|||
Db 1078 GCGCTCTCCCTCCATGACCTCCGCCATGCCCATTTGGAGGAGCAAGCTGTGGCCAAATA 1137
|||
QY 2273 CCAGGAGATCCCTTTTATCATCTGCAATGGGACGAGCGGTACTTGCATGACCTT 2332
|||
Db 1138 CCAGGAGATCCCTTTTATCATCTTGAAGTGGCAGCCAGAGGCGCTCCACTGCACTTT 1197
|||
QY 2333 CACCTCGAGCGTGTCAAGCCAG--CACTAGTGAAGTGGCCCTGCAAGCTGTGGTGTGCG 2391
|||
Db 1198 CACCTCGAGGAGGACAGGCTTGGCTTCAAGAGCTCACCTGCAAGATCTCGGTGCGGC 1257
|||
QY 2392 AGTGGAGGCGAGCGGCGAGCTTCAAGTCACTTCAACATCAACCAAG--GACACAA 2448
|||
Db 1258 AAGTGAAGGGAGGGCGAGATATTTCCAGCTGCATACCACCTCTGCGCAGAGACACCTGTG 1317
|||
QY 2449 GGTTCGTGAGTGTGCTGTGGAGAGTGAAGCGGGGTCCAGCCCTGTGGGCCCA 2508
|||
Db 1318 GTCCTCGGACACTCTGTGCTTGCCTGCGCAGCACTGTCAACCCAGCTGGGACCTT 1377
|||
QY 2509 GTGCTTCAAGATCCCTTCTCATTCGGCAGAGATATTTCCAGCTGGAGCCACCCCT 2568
|||
Db 1378 ATGCTTCAAGATCCCACTGTCCATCCGCGAGAGATGCAACAGCTTAGTGCCCA 1437
|||
QY 2569 GTAGCGGGGTGCCGACTTGGCGGACTTGGCCCGAGAACTCCACTGACAGCACTTCA 2628
|||
Db 1438 ACTCAGCGGGCAATGACTTGGCGGATGTTAGCACAGAGCTCTCTATGACCCGTTACCTGA 1497
|||
QY 2629 GCTTCTTGGCTTCCAAAGCCAGCCCAAGCCATGATCTCAACCTGTGGAGGCGCGC 2688
|||
Db 1498 ATTAATTGGCCAAAGAGCGAGCCCGGTTGTGATCTCTGGAGCTCTGGGAAGCTTGC 1557
|||
QY 2689 ACTTCCCCCAAGCAACCTCAGCGAGCTGGCTGCAGAGTGG 2730
|||
Db 1558 AGCAGGACGATGGGACCTCAAGAGCTGNCAGTGCCTTGG 1599
|||
```

```
Patent No. 5691147
GENERAL INFORMATION:
APPLICANT: Gyuris, Jenő
APPLICANT: Draetta, Giulio
TITLE OF INVENTION: CDK4 Binding Proteins
NUMBER OF SEQUENCES: 95
CORRESPONDENCE ADDRESS:
ADDRESSER: LAHIVE & COCKFIELD
STREET: 60 State Street
CITY: Boston
STATE: MA
COUNTRY: USA
ZIP: 02109
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: ASCII(text)
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/253,155A
FILING DATE: 02-JUN-1994
CLASSIFICATION: 435
ATTORNEY/AGENT INFORMATION:
NAME: Vincent, Matthew P.
REGISTRATION NUMBER: 36,709
REFERENCE/DOCKET NUMBER: MII-028
TELECOMMUNICATION INFORMATION:
TELEPHONE: (617) 227-7400
TELEFAX: (617) 227-5941
INFORMATION FOR SEQ ID NO: 17:
SEQUENCE CHARACTERISTICS:
LENGTH: 771 base pairs
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: cdna
US-08-253-155A-17

Query Match 8.9%; Score 256; DB 1; Length 771;
Best Local Similarity 67.2%; Pred. No. 1.3e-46;
Matches 406; Conservative 1; Mismatches 191; Indels 6; Gaps 3;

QY 1741 AGACCTGTGAGTCCATCGTTAGCTGTGGAGCCCTCGCGTCTGCTCACCGGCCAG 1800
DB 738 AAACAGTATTGAGCCCTCGGTGACCTGTGGACCCACAGCCCTCTGCTGCGGCCCG 679
QY 1801 TCATCTGTGCTATGGACCACTGTGGGAGCCAGCCCTGACAGCTGGAGCCTGCGCTCA 1860
DB 678 TCATCTCTACCATGCCCACTGTGCCGAAGTCAGTGCCTGACTGGATCTTTCAGCTCA 619
QY 1861 AAAGCAGTGTGCGAGGCGAGCTGGAGCAGAGTGTGTCACCTGGGCGAGGAGCGC 1920
DB 618 AGACCCAGCCCAACAGGCGCACTGGG---AGGAGGTGGTACCTGGATGAGGAGCCC 562
QY 1921 CCTCCACCTCTACTACTGCCAGCTGGAGCCAGTGCCTGCTACGCTTTCACCGAGCAG 1980
DB 561 TGAACACACCTGTCTACTGCCAGCTGGAGCCAGGCGCTGTACATCTGCTGGACCAGC 502
QY 1981 TGGGCGGCTTTGCCCTGTGTGGAGAGCCCTACAGCTGGCTGCGCCCAAGCGCTCAAGC 2040
DB 501 TGGGCACTTACGCTGTGTTTCAAGGCGAGTCTTATTCGCGCTCAGCAGTCAAGCGGCTCAGC 442
QY 2041 TGTCTTGTGTCGCGCGGTGGCTGACCTCCTCGAGTACACATCCGGGTCTACTGCG 2100
DB 441 TGGCGGTCTTTCGCCCCCGGCTCTGCACTTCCCTGGAGTACAGACTCCGGGTCTACTGTC 382
QY 2101 TGCATGACCCACAGATGACATCAAGAGGTGGTGCAGCTGGAGAGACGAGCTGGGGGAC 2160
DB 381 TGGAGACAGGCGCTGTAGCACTGAAGAGGTGCTGGAGCTGGAGCGGACTCTCTGGCGGAT 322
QY 2161 AGCTGATCCAGGAGCCAGGCTCTGCACTTCAAGGACAGTACACCAAGCTCGGCTAT 2220
DB 321 ACTTGTGGAGGAGCCGAAACCGCTAATGTTCAAGGACAGTAAACACCAAGCTCGC-CTCT 263

2221 CCATCCAGATGTGCCAGCTCCCTGTGGAAGTAGAAGCTCCTTGTGCTACCTACGAGAGA 2280
DB 262 TCCTCATGACCTCCCCCATGCCATTTGGAGGAGCAAGCTGTGCCAATACGAGAGA 203
QY 2281 TCCCTTTTATCATCTGTGAATGGACGACGCGGTACTTTGCACTGCACCTTTCACCTGG 2340
DB 202 TCCCTTCTATCATATTTGGAGTGGAGCCAGAAGG--CCTCCACTGCACCTTTCACCTGG 145
QY 2341 AGCG 2344
DB 144 AGAG 141

RESULT 10
US-08-808-982-4
Sequence 4, Application US/08080982
Patent No. 5939271
GENERAL INFORMATION:
APPLICANT: Tessier-Lavigne, Marc
APPLICANT: Leonardo, E. David
APPLICANT: Hink, Lindsay
APPLICANT: Masu, Masayuki
APPLICANT: Kazuko, Keino-Masu
TITLE OF INVENTION: Netrin Receptors
NUMBER OF SEQUENCES: 8
CORRESPONDENCE ADDRESS:
ADDRESSER: SCIENCE & TECHNOLOGY LAW GROUP
STREET: 268 BUSH STREET, SUITE 3200
CITY: SAN FRANCISCO
STATE: CALIFORNIA
COUNTRY: USA
ZIP: 94104
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patent In Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/808,982
FILING DATE:
CLASSIFICATION: 530
ATTORNEY/AGENT INFORMATION:
NAME: OSMAN, RICHARD A.
REGISTRATION NUMBER: 36,627
REFERENCE/DOCKET NUMBER: UC96-217
TELECOMMUNICATION INFORMATION:
TELEPHONE: (415) 343-4341
TELEFAX: (415) 343-4342
INFORMATION FOR SEQ ID NO: 4:
SEQUENCE CHARACTERISTICS:
LENGTH: 305 base pairs
TYPE: nucleic acid
STRANDEDNESS: double
TOPOLOGY: linear
MOLECULE TYPE: cdna
US-08-808-982-4

Query Match 4.1%; Score 119; DB 2; Length 305;
Best Local Similarity 68.1%; Pred. No. 5.6e-17;
Matches 209; Conservative 0; Mismatches 95; Indels 3; Gaps 3;

QY 1906 TGGGCGAGGAGCGCCCTCCACCTCTACTTGCAGCTGGAGGCGAGTGGCTGTGCTACG 1965
DB 1 TGGATGAGGAGACCTTGAACACACACCTGTGCTACTG-CAGCTGGAGCCAGGGCTG-TACA 58
QY 1966 TTTTACCGAGCAGCTGGGCGCGCTTTGCCCTGTGGAGAGGCCCTTCAGCGTGGTGGCGG 2025
DB 59 TCTGCTGACACAGCTGGGCACTACGTTTTCAGGGCGAGTCTTATTCGCGCTCAGCAG 118
QY 2026 CCAAGCGCTCAAGCTGTCTTGTGCGCGGTGGCTTGCACCTCCCTCGAGTACACA 2085
DB 119 TCAAGCGGTCCAGTGGCGGT-TTCGCCCCCGCCCTCTGCACTCCCTCGGAGTACAGCC 177
```

Qy	2086	TCGGGTCTACTGCTGCATGACACACCACGATGCACCTCAAGGAGGTGGTCAGCTGGAGA	2145
Db	178	TCCGGGTCTACTGCTTGGAGACACGCTGTAGCACTCAAGGAGGTCTGGAGCTGGAGC	237
Qy	2146	AGCAGCTGGGGGGACAGTGTATCAGGAGCCACGGGTCTCTGCATTTCAAGACAGATTACC	2205
Db	238	GGACTCTGGCGCGATACTTGGTGGAGAGCCGAACCGCTAAATGTTCAAGGACAGTTACC	297
Qy	2206	ACAACCT	2212
Db	298	ACAACCT	304

```

1
2
3
4
5
6
7
8
9
10
11
12
13
14
15
16
17
18
19
20
21
22
23
24
25
26
27
28
29
30
31
32
33
34
35
36
37
38
39
40
41
42
43
44
45
46
47
48
49
50
51
52
53
54
55
56
57
58
59
60
61
62
63
64
65
66
67
68
69
70
71
72
73
74
75
76
77
78
79
80
81
82
83
84
85
86
87
88
89
90
91
92
93
94
95
96
97
98
99
100
101
102
103
104
105
106
107
108
109
110
111
112
113
114
115
116
117
118
119
120
121
122
123
124
125
126
127
128
129
130
131
132
133
134
135
136
137
138
139
140
141
142
143
144
145
146
147
148
149
150
151
152
153
154
155
156
157
158
159
160
161
162
163
164
165
166
167
168
169
170
171
172
173
174
175
176
177
178
179
180
181
182
183
184
185
186
187
188
189
190
191
192
193
194
195
196
197
198
199
200
201
202
203
204
205
206
207
208
209
210
211
212
213
214
215
216
217
218
219
220
221
222
223
224
225
226
227
228
229
230
231
232
233
234
235
236
237
238
239
240
241
242
243
244
245
246
247
248
249
250
251
252
253
254
255
256
257
258
259
260
261
262
263
264
265
266
267
268
269
270
271
272
273
274
275
276
277
278
279
280
281
282
283
284
285
286
287
288
289
290
291
292
293
294
295
296
297
298
299
300
301
302
303
304
305
306
307
308
309
310
311
312
313
314
315
316
317
318
319
320
321
322
323
324
325
326
327
328
329
330
331
332
333
334
335
336
337
338
339
340
341
342
343
344
345
346
347
348
349
350
351
352
353
354
355
356
357
358
359
360
361
362
363
364
365
366
367
368
369
370
371
372
373
374
375
376
377
378
379
380
381
382
383
384
385
386
387
388
389
390
391
392
393
394
395
396
397
398
399
400
401
402
403
404
405
406
407
408
409
410
411
412
413
414
415
416
417
418
419
420
421
422
423
424
425
426
427
428
429
430
431
432
433
434
435
436
437
438
439
440
441
442
443
444
445
446
447
448
449
450
451
452
453
454
455
456
457
458
459
460
461
462
463
464
465
466
467
468
469
470
471
472
473
474
475
476
477
478
479
480
481
482
483
484
485
486
487
488
489
490
491
492
493
494
495
496
497
498
499
500
501
502
503
504
505
506
507
508
509
510
511
512
513
514
515
516
517
518
519
520
521
522
523
524
525
526
527
528
529
530
531
532
533
534
535
536
537
538
539
540
541
542
543
544
545
546
547
548
549
550
551
552
553
554
555
556
557
558
559
560
561
562
563
564
565
566
567
568
569
570
571
572
573
574
575
576
577
578
579
580
581
582
583
584
585
586
587
588
589
590
591
592
593
594
595
596
597
598
599
600
601
602
603
604
605
606
607
608
609
610
611
612
613
614
615
616
617
618
619
620
621
622
623
624
625
626
627
628
629
630
631
632
633
634
635
636
637
638
639
640
641
642
643
644
645
646
647
648
649
650
651
652
653
654
655
656
657
658
659
660
661
662
663
664
665
666
667
668
669
670
671
672
673
674
675
676
677
678
679
680
681
682
683
684
685
686
687
688
689
690
691
692
693
694
695
696
697
698
699
700
701
702
703
704
705
706
707
708
709
710
711
712
713
714
715
716
717
718
719
720
721
722
723
724
725
726
727
728
729
730
731
732
733
734
735
736
737
738
739
740
741
742
743
744
745
746
747
748
749
750
751
752
753
754
755
756
757
758
759
760
761
762
763
764
765
766
767
768
769
770
771
772
773
774
775
776
777
778
779
780
781
782
783
784
785
786
787
788
789
790
791
792
793
794
795
796
797
798
799
800
801
802
803
804
805
806
807
808
809
810
811
812
813
814
815
816
817
818
819
820
821
822
823
824
825
826
827
828
829
830
831
832
833
834
835
836
837
838
839
840
841
842
843
844
845
846
847
848
849
850
851
852
853
854
855
856
857
858
859
860
861
862
863
864
865
866
867
868
869
870
871
872
873
874
875
876
877
878
879
880
881
882
883
884
885
886
887
888
889
890
891
892
893
894
895
896
897
898
899
900
901
902
903
904
905
906
907
908
909
910
911
912
913
914
915
916
917
918
919
920
921
922
923
924
925
926
927
928
929
930
931
932
933
934
935
936
937
938
939
940
941
942
943
944
945
946
947
948
949
950
951
952
953
954
955
956
957
958
959
960
961
962
963
964
965
966
967
968
969
970
971
972
973
974
975
976
977
978
979
980
981
982
983
984
985
986
987
988
989
990
991
992
993
994
995
996
997
998
999
1000
1001
1002
1003
1004
1005
1006
1007
1008
1009
1010
1011
1012
1013
1014
1015
1016
1017
1018
1019
1020
1021
1022
1023
1024
1025
1026
1027
1028
1029
1030
1031
1032
1033
1034
1035
1036
1037
1038
1039
1040

```

Db	119	TC AAGCGGCTCCAGCTGGCCGT -TTCCGCCCCCGCCCTCTGCACCTCCCTGGAGTACAGCC	177
Qy	2086	TC CGGGTCTACTGCCTGCATGCACACCACGATGCACCTCAAGGAGGTGGTGCACGTGAGA	2145
Db	178	TC CGGGTCTACTGCCTGGAGGACACGCCCTGTAGCACTGAAGGAGGTGCTGGAGCTGGAGC	237
Qy	2146	AGCAGCTGGGGGGACAGCTGATCCAGGAGCCACGGGTCCCTGCATTC AAGACACAGTTTACC	2205
Db	238	GGACTCTGGCGCGATCTTGGTGGAGAGCCGAAACCGCTAAATGTTCAAGACACAGTTTACC	297
Qy	2206	ACAACCT	2212
Db	298	ACAACCT	304

RESULT 12
 US-08-483-533-38
 ; Sequence 38, Application US/08483533
 ; Patent No. 6172047
 ; GENERAL INFORMATION:
 ; APPLICANT: Roizman, Bernard
 ; APPLICANT: Chou, Joany
 ; TITLE OF INVENTION: Method for Treating Tumorigenic
 ; TITLE OF INVENTION: Diseases
 ; NUMBER OF SEQUENCES: 43
 ; CORRESPONDENCE ADDRESS:
 ; ADDRESSEE: Marshall, O'Toole, Gerstein, Murray & Borun
 ; STREET: 6300 Sears Tower, 233 South Wacker Drive
 ; CITY: Chicago
 ; STATE: Illinois
 ; COUNTRY: United States of America
 ; ZIP: 60606-6402
 ; COMPUTER READABLE FORM:
 ; MEDIUM TYPE: Floppy disk
 ; COMPUTER: IBM PC compatible
 ; OPERATING SYSTEM: PC-DOS/MS-DOS
 ; SOFTWARE: Patent In Release #1.0, Version #1.25
 ; CURRENT APPLICATION DATA:
 ; APPLICATION NUMBER: US/08/483,533
 ; FILING DATE: 07-MAR-95
 ; CLASSIFICATION: 514
 ; PRIOR APPLICATION NUMBER: 08/419,853
 ; FILING DATE: 11-APR-95
 ; PRIOR APPLICATION NUMBER: 07/861,233
 ; FILING DATE: 31-MAR-92
 ; ATTORNEY/AGENT INFORMATION:
 ; NAME: Zeller, James P.
 ; REGISTRATION NUMBER: 28,491
 ; REFERENCE/DOCKET NUMBER: 28097/32742
 ; TELECOMMUNICATION INFORMATION:
 ; TELEPHONE: 312/474-6300
 ; TELEFAX: 312/474-0448
 ; TELEFAX: 25-3856
 ; INFORMATION FOR SEQ ID NO: 38:
 ; SEQUENCE CHARACTERISTICS:
 ; LENGTH: 1280 base pairs
 ; TYPE: nucleic acid
 ; STRANDEDNESS: single
 ; TOPOLOGY: linear
 ; MOLECULE TYPE: DNA (genomic)
 US-08-483-533-38

Query Match	1.8%	Score 51.8;	DB 3;	Length 1280;
Best Local Similarity	51.0%;	Pred. No. 0.034;		
Matches 122;	Conservative 0;	Mismatches 117;	Indels 0;	Gaps 0

QY	6	GGGCTCGGGGTGAGGGCTAAAGCCGCCCTCCCGCCCGGGGGCCCGCGCCGCCGCCG	65
Db	555	GCGGCGGGGAGGGGGCGCGAAGCCCGCCGACACCCCGACACCCCGACCCCGCG	614
QY	66	CCCGCTCCCGCCCGGGGGCATGGCCGTCGCGCCCGGGCTGTGGCCAGCGCTCTGGGC	125

Db 615 ACCCCGCGACCCCGCGACCCCGCGACCCCGCGCGGTGCGCTTCTGCGCCCAAGTC 674
QY 126 ATAGTCCTCGCGCTTGGCTCGCGCGCTCGGGTGCACAGAGTGCACCGTGGCCCAAC 185
Db 675 CGGTGCGCCACCTGGTGTCTGGGCTCGCGCGCGCGCTCGCGCGCGCTCGTGG 734
QY 186 CAGTGTCTGGTGCACCGGACCTGCTTCCCACTTCTGCTGGAGCCGAGGATGT 244
Db 735 GCCCGGAGCGGCGCGCGCTCGGTTCCGCGCGCGGTGCGGAGCGCGGT 793

RESULT 13

US-09-283-471A-38

; Sequence 38, Application US/09283471A

; Patent No. 6340673

; GENERAL INFORMATION:

; APPLICANT: Roizman, Bernard

; APPLICANT: Chou, Jony

; TITLE OF INVENTION: Method For Treating Tumorigenic Diseases

; NUMBER OF SEQUENCES: 43

; CORRESPONDENCE ADDRESS:

; ADDRESSEE: Marshall, O'Toole, Gerstein, Murray & Borun

; STREET: 6300 Sears Tower, 233 South Wacker Drive

; CITY: Chicago

; STATE: Illinois

; COUNTRY: United States of America

; ZIP: 60606-6402

; COMPUTER READABLE FORM:

; MEDIUM TYPE: Floppy disk

; COMPUTER: IBM PC compatible

; OPERATING SYSTEM: PC-DOS/MS-DOS

; SOFTWARE: Patent in Release #1.0, Version #1.25

; CURRENT APPLICATION DATA:

; APPLICATION NUMBER: US/09/283,471A

; FILING DATE: 04-APR-1999

; CLASSIFICATION: 514

; PRIOR APPLICATION DATA:

; APPLICATION NUMBER: 07/861,233

; FILING DATE: 31-MAR-1992

; PRIOR APPLICATION DATA:

; APPLICATION NUMBER: 08/419,853

; FILING DATE: 11-APR-1995

; PRIOR APPLICATION DATA:

; APPLICATION NUMBER: 08/483,533

; FILING DATE: 07-JUN-1995

; ATTORNEY/AGENT INFORMATION:

; NAME: Zeller, James P.

; REGISTRATION NUMBER: 28,491

; REFERENCE/DOCKET NUMBER: 27373/32742A

; TELECOMMUNICATION INFORMATION:

; TELEPHONE: 312/474-6300

; TELEFAX: 312/474-0448

; TELEX: 25-3856

; INFORMATION FOR SEQ ID NO: 38:

; SEQUENCE CHARACTERISTICS:

; LENGTH: 1280 base pairs

; TYPE: nucleic acid

; STRANDEDNESS: single

; TOPOLOGY: linear

; MOLECULE TYPE: DNA (genomic)

US-09-283-471A-38

Query Match 1.8%; Score 51.8; DB 4; Length 1280;

Best Local Similarity 51.0%; Pred. No. 0.034;

Matches 122; Conservative 0; Mismatches 117; Indels 0; Gaps 0;

QY 6 GCGCTCGGCTGAGCGCTAAGCGCCCTCGCGCGGCGCGCGCGCGCGCGCGCG 65

Db 555 GCGGCGGCGGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 614

QY 66 CCGGCTGCG 125

Db 615 ACCCCGCGACCCCGCGACCCCGCGACCCCGCGCGGTGCGCTTCTGCGCCCAAGTC 674
QY 126 ATAGTCCTCGCGCTTGGCTCGCGCGCTCGGGTGCACAGAGTGCACCGTGGCCCAAC 185
Db 675 CGGTGCGCCACCTGGTGTCTGGGCTCGCGCGCGCGCTCGCGCGCGCTCGTGG 734
QY 186 CAGTGTCTGGTGCACCGGACCTGCTTCCCACTTCTGCTGGAGCCGAGGATGT 244
Db 735 GCCCGGAGCGGCGCGCGCTCGGTTCCGCGCGCGGTGCGGAGCGCGGT 793

RESULT 14

US-09-252-991A-7947

; Sequence 7947, Application US/09252991A

; Patent No. 6551795

; GENERAL INFORMATION:

; APPLICANT: Marc J. Rubenfield et al.

; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO PSEUDOMONAS

; TITLE OF INVENTION: AERUGINOSA FOR DIAGNOSTICS AND THERAPEUTICS

; FILE REFERENCE: 107196.136

; CURRENT APPLICATION NUMBER: US/09/252,991A

; CURRENT FILING DATE: 1999-02-18

; PRIOR APPLICATION NUMBER: US 60/074,788

; PRIOR FILING DATE: 1998-02-18

; PRIOR APPLICATION NUMBER: US 60/094,190

; PRIOR FILING DATE: 1998-07-27

; NUMBER OF SEQ ID NOS: 33142

; SEQ ID NO 7947

; TYPE: DNA

; LENGTH: 699

; ORGANISM: Pseudomonas aeruginosa

US-09-252-991A-7947

Query Match 1.8%; Score 51.2; DB 4; Length 699;

Best Local Similarity 47.5%; Pred. No. 0.036;

Matches 152; Conservative 0; Mismatches 168; Indels 0; Gaps 0;

QY 1876 AGGCGAGCTGGGAGCAGGATGTGTGCACCTGGGCGAGGAGGCGCCCTCCACCTTACT 1935

Db 168 ACAGCTGTACGACGACGAGGTGCTCGAGGCCAAGTGTCTGTCGGGCGCATTCGGCT 227

QY 1936 ACTGCCAGCTGGAGGCCAGTGTCTGTCTTACCCAGAGCTGGGCGCGCTTTGCC 1995

Db 228 ACGACCTGGTGTGCGCAGCAGCTTCTCTGCCAATCTTCAAGGCGGAGGTTC 287

QY 1996 TGTGGGAGAGGCGCTCAGCTGGTGGCGCCAGCGCTCAGCTGCTTCTGTTGCGC 2055

Db 288 AGCCGCTGGACAAGAGCAAGTGTCCGAACTGGAAGAACTCAACCGGCGCTCTCAAGG 347

QY 2056 CGTGGCTTGCACCTCCCTCGAGTACAACATCCGGGTCTACTGCTGCTGATGACACCCACG 2115

Db 348 TGCTGCGCGGAGGACCCCGGCAACCGCTACGTGATGCCCTACATGTGGGCGACCAACG 407

QY 2116 ATGCACTCAAGGAGGTGGTGTGAGTGGAGAGCAGCTGGGGGACAGCTGATCCAGGAGC 2175

Db 408 GCATCGCTACAACCTCGACAAGTGTCCGCGCGGTGCTCGGCGAGCATGCGCGCTGACT 467

QY 2176 CACGGTCTGTCACCTTCAAG 2195

Db 468 CTGGGAGCTGTGTTCAG 487

RESULT 15

US-09-252-991A-7809

; Sequence 7809, Application US/09252991A

; Patent No. 6551795

; GENERAL INFORMATION:

; APPLICANT: Marc J. Rubenfield et al.

; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO PSEUDOMONAS

; TITLE OF INVENTION: AERUGINOSA FOR DIAGNOSTICS AND THERAPEUTICS

; FILE REFERENCE: 107196.136

; CURRENT APPLICATION NUMBER: US/09/252,991A

; CURRENT FILING DATE: 1999-02-18

```
; PRIOR APPLICATION NUMBER: US 60/074,788
; PRIOR FILING DATE: 1998-02-18
; PRIOR APPLICATION NUMBER: US 60/094,190
; PRIOR FILING DATE: 1998-07-27
; NUMBER OF SEQ ID NOS: 33142
; SEQ ID NO 7809
; LENGTH: 1302
; TYPE: DNA
; ORGANISM: Pseudomonas aeruginosa
US-09-252-991A-7809
```

```
Query Match      1.8%; Score 51.2; DB 4; Length 1302;
Best Local Similarity 47.5%; Pred.No.0.046;
Matches 152; Conservative 0; Mismatches 168; Indels 0; Gaps 0;

QY 1876 AGGGCAGCTGGGAGCAGGATGCTGCACCTGGCGGAGGAGGCGCCCTCCACCTCTACT 1935
Db 368 AGACGCTCTACGACACGCAACGAGGTGCTGAGGCCAAGCTGTCTCCGGGCATTCCGGCT 427

QY 1936 ACTGCCAGCTGGAGGCCAGTGCCTGCTAGCTCTTACCGAGCAGCTGGGCGGCTTTGCCG 1995
Db 428 ACGACCTGCTGTGTCGCCAGCGACAGCTTCTGCCCCAACTACCTGAAGGCGGAGGTGTTCC 487

QY 1996 TGGTGGGAGAGGCCCTCAGCGTGGCTGCCGCCAAGCGCCTCAAGCTGCTTCTTTTGGCG 2055
Db 488 AGCCGCTGGACAAGAGCAAGCTGCCGAAGTGGAAAGAACCTCAACCCGGCCCTGCTCAAG 547

QY 2056 CGGTGGCCTGACCTCCCTCGAGTACAACATCGGGTCTACTGCCCTGCATGACCCACG 2115
Db 548 TGCTGCCCGGCAAGGACCCCGGCAACCGTACGTGATGCCCTACATGTGGGGCACCAACG 607

QY 2116 ATGCACTCAAGGAGTGTGTCAGCTGGAGAAGCAGCTGGGGGGACAGCTGATCCAGGAGC 2175
Db 608 GCATCGCCTACAACTCGACAAGTCCGCGCGGTGCTCGGCGACGATGCGCGCTGGACT 667

QY 2176 CACGGGTCTCGACTTCAAG 2195
Db 668 CTTGGACCTGGTGTTCAG 687
```

Search completed: October 5, 2004, 20:46:56
Job time : 210 secs

This Page Blank (uspto)

1	2881	100.0	2881	11	US-09-970-944-1	Sequence 1, Appl
2	2676.4	92.9	2752	13	US-09-918-779-1	Sequence 1, Appl
3	2676.4	92.9	2752	17	US-10-624-932-1	Sequence 1, Appl
4	2402.6	83.4	3580	17	US-10-311-623-13	Sequence 13, Appl
5	2235.8	77.6	3014	10	US-09-933-261-1	Sequence 1, Appl
6	2235.8	77.6	3014	15	US-10-256-702-1	Sequence 1, Appl
7	2189.4	76.0	2697	15	US-10-240-154-15	Sequence 15, Appl
8	1570.4	54.5	1787	10	US-09-933-261-2	Sequence 2, Appl
9	1570.4	54.5	1787	15	US-10-256-702-2	Sequence 2, Appl
10	1193.6	41.4	1321	13	US-10-496-115-365	Sequence 365, App
11	1147.8	39.8	1282	9	US-09-833-381-1806	Sequence 1806, Ap
12	889	30.9	2860	13	US-10-087-684-1	Sequence 1, Appl
13	889	30.9	2860	13	US-09-918-779-1	Sequence 1, Appl
14	887.4	30.8	2860	13	US-10-087-684-3	Sequence 3, Appl

	Query Match	100.0%;	Score 2881;	DB 11;	Length 2881;	
	Best Local Similarity	100.0%;	Pred. No. 0;			
	Matches 2881;	Conservative	0;	Mismatches	0;	Indels
						Gaps
						0;
QY	1	AGCTGGGGCTCCGGGCTGAGGCGCTAAAGCGCGCCCTCCCGCGCGGGGCCCCCGCGCCG	60			
Db	1	AGCTGGGGCTCCGGGCTGAGGCGCTAAAGCGCGCCCTCCCGCGCGGGGCCCCCGCGCCG	60			
QY	61	GCCCGCGCGCCCTGCCCGCGCGCGCCATGGCGTCCGCGCCCGGACCTGTGCACGAGCGTCC	120			
Db	61	GCCCGCGCGCCCTGCCCGCGCGCGCCATGGCGTCCGCGCCCGGACCTGTGCACGAGCGTCC	120			
QY	121	TGGGATAGTCTCTCGCGGCTTGGCTCCGGGCTCGGGTGCACGAGAGTGCACACGTGG	180			

Db 121 TGGGCAATAGTCTCGCGGCTTGGCTCGGGCTCGGGTCCCGACGAGTGGCCACCGTGG 180
Qy 181 CCAACCCAGTGTCTGGTGCCAAACCGGACCTCTCTCCCACTTCTCTGGTGAGCCCGAGG 240
Db 181 CCAACCCAGTGTCTGGTGCCAAACCGGACCTCTCTCCCACTTCTCTGGTGAGCCCGAGG 240
Qy 241 ATGTGTACATCTCTAAGAACAGCCAGTGTCTCTGTGTGCAAGCCGCTGCGCCGACGC 300
Db 241 ATGTGTACATCTCTAAGAACAGCCAGTGTCTCTGTGTGCAAGCCGCTGCGCCGACGC 300
Qy 301 AGATCTTCTCAAGTGTCAACGGGAGTGGTGCGCCAGGTGCACACAGTGTGATCGAGCGCA 360
Db 301 AGATCTTCTCAAGTGTCAACGGGAGTGGTGCGCCAGGTGCACACAGTGTGATCGAGCGCA 360
Qy 361 GCACAGACGGGAGCGTGTGTGAGCCGACCATGAGGTCGCGATTAATGTCTCAAGSCAGC 420
Db 361 GCACAGACGGGAGCGTGTGTGAGCCGACCATGAGGTCGCGATTAATGTCTCAAGSCAGC 420
Qy 421 AGGTGAGAGAGTGTTCGGGCTGGAGGAATACGTGTGCCAGTGCCTGGGATGAGCTCCT 480
Db 421 AGGTGAGAGAGTGTTCGGGCTGGAGGAATACGTGTGCCAGTGCCTGGGATGAGCTCCT 480
Qy 481 CGGGACCAACCAAGAGTCAGAGGCTTACATCCGATAGCCAGATTGCGCAAGAACTTCG 540
Db 481 CGGGACCAACCAAGAGTCAGAGGCTTACATCCGATAGCCAGATTGCGCAAGAACTTCG 540
Qy 541 AGCAGAGCGGTGTGCGCAAGGAGTGTCTCTGAGCAGGGCATGTGTGCTGCCCTGCCGTC 600
Db 541 AGCAGAGCGGTGTGCGCAAGGAGTGTCTCTGAGCAGGGCATGTGTGCTGCCCTGCCGTC 600
Qy 601 CACCGAGAGCATCCCTCCAGCCGAGGTGGAGTGGCTCCGGACAGAGGACCTGTGTGAGC 660
Db 601 CACCGAGAGCATCCCTCCAGCCGAGGTGGAGTGGCTCCGGACAGAGGACCTGTGTGAGC 660
Qy 661 CGTCCCTGGACCCCAATATATACATCAGCGGAGCACAGCTGTGTGTCGACAGGCC 720
Db 661 CGTCCCTGGACCCCAATATATACATCAGCGGAGCACAGCTGTGTGTCGACAGGCC 720
Qy 721 GCCTTGTCTGACCGGCCAATACATCTGCTGTGCGCAAGAACATCTGTGGCAGCTGCGCGCA 780
Db 721 GCCTTGTCTGACCGGCCAATACATCTGCTGTGCGCAAGAACATCTGTGGCAGCTGCGCGCA 780
Qy 781 GCGCTCCCTGTGTCTGTCTGTCTGTCTGTCTGTCTGTCTGTCTGTCTGTCTGTCTGTCT 840
Db 781 GCGCTCCCTGTGTCTGTCTGTCTGTCTGTCTGTCTGTCTGTCTGTCTGTCTGTCTGTCT 840
Qy 841 CCGTCTGACGCGCAGTGTGGCGGCTGTGGAGAAACGAGCCGAGCTGACCAACC 900
Db 841 CCGTCTGACGCGCAGTGTGGCGGCTGTGGAGAAACGAGCCGAGCTGACCAACC 900
Qy 901 CGGCGCTCTCAACCGGGGGCGCTTCTGTGAGGGGCGAGATGTCCATGACCGCACCGTCT 960
Db 901 CGGCGCTCTCTCAACCGGGGGCGCTTCTGTGAGGGGCGAGATGTCCATGACCGCACCGTCT 960
Qy 961 CCTCTCTGTGTCTGTGTGAGCGGAGCTGGAGCCGCTGAGCAAGTGTGCGCCTGTG 1020
Db 961 CCTCTCTGTGTCTGTGTGAGCGGAGCTGGAGCCGCTGAGCAAGTGTGCGCCTGTG 1020
Qy 1021 GGCTGAGTGCACCCCACTGCGGAGCGGTGAGTGTCTGTGACCCAGCACCCCGCAACGAG 1080
Db 1021 GGCTGAGTGCACCCCACTGCGGAGCGGTGAGTGTCTGTGACCCAGCACCCCGCAACGAG 1080
Qy 1081 GGGAGAGTGCAGGGACCTGACCTGACACCCGCACTGTACAGTGTGCTGTGTATC 1140
Db 1081 GGGAGAGTGCAGGGACCTGACCTGACACCCGCACTGTACAGTGTGCTGTGTATC 1140
Qy 1141 ACAGTGTCTTGGCCCTGAGGACGTGGCCCTCTATGTGGGCTCATCGCGTGGCGCTCT 1200
Db 1141 ACAGTGTCTTGGCCCTGAGGACGTGGCCCTCTATGTGGGCTCATCGCGTGGCGCTCT 1200
Qy 1201 GCGTGTCTCTGTCTGTCTGTCTGTCTGTCTGTCTGTCTGTCTGTCTGTCTGTCTGTCTGT 1260
Db 1201 GCGTGTCTCTGTCTGTCTGTCTGTCTGTCTGTCTGTCTGTCTGTCTGTCTGTCTGTCTGT 1260

Qy 1261 ACTCAGATGTGGCTGACTGTCTCACTTCTCACTCAGGTTTCCAGCCCGTTCAGCATCAAGC 1320
Db 1261 ACTCAGATGTGGCTGACTGTCTCACTTCTCACTCAGGTTTCCAGCCCGTTCAGCATCAAGC 1320
Qy 1321 CCAGCAAGAGCAGCAACCCCATCTGTCTACCATCAGCCGAGCTTCAGCACCAACA 1380
Db 1321 CCAGCAAGAGCAGCAACCCCATCTGTCTACCATCAGCCGAGCTTCAGCACCAACA 1380
Qy 1381 CCTACAGGGCAGTCTGTCTCCCGGAGGATGGGCCAGCCCGAGTTTCCAGCTCACCA 1440
Db 1381 CCTACAGGGCAGTCTGTCTCCCGGAGGATGGGCCAGCCCGAGTTTCCAGCTCACCA 1440
Qy 1441 ATGGGCACTGTCTCAGCCCTGTGGTGGCGGCGCCACACACTGTCACCAACAGCTTCCCA 1500
Db 1441 ATGGGCACTGTCTCAGCCCTGTGGTGGCGGCGCCACACACTGTCACCAACAGCTTCCCA 1500
Qy 1501 CCTCTGAGCCGAGGAGTTCGTCTCCCGCTCTCCACCCAGAACTACTTCCGCTCCCTGC 1560
Db 1501 CCTCTGAGCCGAGGAGTTCGTCTCCCGCTCTCCACCCAGAACTACTTCCGCTCCCTGC 1560
Qy 1561 CCGGAGGACACGACATGACCTATGGAGCTTCAACTTCTCGGGGCGGCTGATGA 1620
Db 1561 CCGGAGGACACGACATGACCTATGGAGCTTCAACTTCTCGGGGCGGCTGATGA 1620
Qy 1621 TCCCTAATACAGTATCAGCTTCTCATCCCGCATGATGCCATACCCCGAGGAAATCT 1680
Db 1621 TCCCTAATACAGTATCAGCTTCTCATCCCGCATGATGCCATACCCCGAGGAAATCT 1680
Qy 1681 ATGAGATCTACCTCAGCTGCAAGCCGAGACGTGAGTTGGCTTGGCTGCTC 1740
Db 1681 ATGAGATCTACCTCAGCTGCAAGCCGAGACGTGAGTTGGCTTGGCTGCTC 1740
Qy 1741 AGACCTCTGAGTCCCATCTGTGTGAGCCCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1800
Db 1741 AGACCTCTGAGTCCCATCTGTGTGAGCCCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1800
Qy 1801 TCATCTCTGTGTGAGACCACTGTGGGAGCCAGCCCTGACAGCTGAGGCTTCGCTGCTCA 1860
Db 1801 TCATCTCTGTGTGAGACCACTGTGGGAGCCAGCCCTGACAGCTGAGGCTTCGCTGCTCA 1860
Qy 1861 AAAAGCAGTCTGTCGAGGGCAGCTGGGAGCAGGATGTGTGACCTTGGGCGAGGAGCGC 1920
Db 1861 AAAAGCAGTCTGTCGAGGGCAGCTGGGAGCAGGATGTGTGACCTTGGGCGAGGAGCGC 1920
Qy 1921 CTTCCCACTCTACTACTGCGAGCTGGAGCCAGTCCCTGCTACGCTTTCACCGAGCAGC 1980
Db 1921 CTTCCCACTCTACTACTGCGAGCTGGAGCCAGTCCCTGCTACGCTTTCACCGAGCAGC 1980
Qy 1981 TGGGCGCTTGTGGCTGTGGGAGAGCCCTCAGCGTGGCTGCGCCCAAGCGCTCAAGC 2040
Db 1981 TGGGCGCTTGTGGCTGTGGGAGAGCCCTCAGCGTGGCTGCGCCCAAGCGCTCAAGC 2040
Qy 2041 TGCTTCTGTTTGGCGCGTGTGCACTTCCCTCGAGTACAACTCCGGGTCTACTGCC 2100
Db 2041 TGCTTCTGTTTGGCGCGTGTGCACTTCCCTCGAGTACAACTCCGGGTCTACTGCC 2100
Qy 2101 TGCAATGACACCCAGATGCACTCAAGAGGTGTGAGCTGGAGAGACAGCTGCGGGGAC 2160
Db 2101 TGCAATGACACCCAGATGCACTCAAGAGGTGTGAGCTGGAGAGACAGCTGCGGGGAC 2160
Qy 2161 AGCTGATCCAGGAGCAGCGGTCTGCACTTCAAGGACAGTTACCAACCTGCGCTAT 2220
Db 2161 AGCTGATCCAGGAGCAGCGGTCTGCACTTCAAGGACAGTTACCAACCTGCGCTAT 2220
Qy 2221 CCATCCAGATGTGCCAGCTCCCTGTGGAAGTAACTCTTGTGCTACAGTACAGAGA 2280
Db 2221 CCATCCAGATGTGCCAGCTCCCTGTGGAAGTAACTCTTGTGCTACAGTACAGAGA 2280
Qy 2281 TCCCTTTTATCATCTCTGAATGGCAGCGGATCTTGCATGCACTGCACTTCACTTGG 2340
Db 2281 TCCCTTTTATCATCTCTGAATGGCAGCGGATCTTGCATGCACTGCACTTCACTTGG 2340

Qy	2341	AGCGTGTGAGCCCGACGACTAGTGACTGGGCTGCAAGCTGTGGGTGTGGCAGGTGGAGG	2400
Db	2341	AGCGTGTGAGCCCGACGACTAGTGACTGGGCTGCAAGCTGTGGGTGTGGCAGGTGGAGG	2400
		AGCGTGTGAGCCCGACGACTAGTGACTGGGCTGCAAGCTGTGGGTGTGGCAGGTGGAGG	2400
Qy	2401	GCAGCGGCAGAGCTTCAGCATCAACTTCAACATCACCAAGCACACAGGTTTCTGTGACG	2460
Db	2401	GCAGCGGCAGAGCTTCAGCATCAACTTCAACATCACCAAGCACACAGGTTTCTGTGACG	2460
		GCAGCGGCAGAGCTTCAGCATCAACTTCAACATCACCAAGCACACAGGTTTCTGTGACG	2460
Qy	2461	TGCTGGCTCTCGAGAGTGAAGCGGGGTCCACAGCCCTGTGTGGGCCCCAGTGCCTTCAAGA	2520
Db	2461	TGCTGGCTCTCGAGAGTGAAGCGGGGTCCACAGCCCTGTGTGGGCCCCAGTGCCTTCAAGA	2520
		TGCTGGCTCTCGAGAGTGAAGCGGGGTCCACAGCCCTGTGTGGGCCCCAGTGCCTTCAAGA	2520
Qy	2521	TCCCTTTCCTCATTTGGGAGAGATAAATTTCCAGCTTGGACCCACCTGTAGGCGGGGTG	2580
Db	2521	TCCCTTTCCTCATTTGGGAGAGATAAATTTCCAGCTTGGACCCACCTGTAGGCGGGGTG	2580
		TCCCTTTCCTCATTTGGGAGAGATAAATTTCCAGCTTGGACCCACCTGTAGGCGGGGTG	2580
Qy	2581	CCGACTGGCGGACTCTGGCCAGAACTCACCTGGACAGGCATCTCAGCTTCTTTGCTCT	2640
Db	2581	CCGACTGGCGGACTCTGGCCAGAACTCACCTGGACAGGCATCTCAGCTTCTTTGCTCT	2640
		CCGACTGGCGGACTCTGGCCAGAACTCACCTGGACAGGCATCTCAGCTTCTTTGCTCT	2640
Qy	2641	CCAAGCCAGCCCAACAGCCATGATCTCAACCTGTGGAGGGCGGCACCTTCCCCAAG	2700
Db	2641	CCAAGCCAGCCCAACAGCCATGATCTCAACCTGTGGAGGGCGGCACCTTCCCCAAG	2700
		CCAAGCCAGCCCAACAGCCATGATCTCAACCTGTGGAGGGCGGCACCTTCCCCAAG	2700
Qy	2701	GCAACTCAGCCAGCTGGCTGCAGCACTGGCTGGACTGGGCACAGACGCTGGCTCT	2760
Db	2701	GCAACTCAGCCAGCTGGCTGCAGCACTGGCTGGACTGGGCACAGACGCTGGCTCT	2760
		GCAACTCAGCCAGCTGGCTGCAGCACTGGCTGGACTGGGCACAGACGCTGGCTCT	2760
Qy	2761	TCCAGTGTGGAGGCTGAGTGTGAGCGCGGCCAGGCCGACACCTTACACTCTCACAG	2820
Db	2761	TCCAGTGTGGAGGCTGAGTGTGAGCGCGGCCAGGCCGACACCTTACACTCTCACAG	2820
		TCCAGTGTGGAGGCTGAGTGTGAGCGCGGCCAGGCCGACACCTTACACTCTCACAG	2820
Qy	2821	CTTTGGCACCCACCAAGGACAGGACGAGCGCGGACAGGGGCCCTTCCCCACAGGGGAG	2880
Db	2821	CTTTGGCACCCACCAAGGACAGGACGAGCGCGGACAGGGGCCCTTCCCCACAGGGGAG	2880
		CTTTGGCACCCACCAAGGACAGGACGAGCGCGGACAGGGGCCCTTCCCCACAGGGGAG	2880
Qy	2881	A 2881	
Db	2881	A 2881	
		A 2881	

RESULT 2

```

US-09-918-779--1
: Sequence 1, Application US/09918779
: Publication No. US20030064369A1
: GENERAL INFORMATION:
: APPLICANT: Taupier, Raymond
: APPLICANT: Padigaru, Muralidhara
: APPLICANT: Rastelli, Luca
: APPLICANT: Spaderna, Steven
: APPLICANT: Shimkets, Richard
: APPLICANT: Zerhusen, Bryan
: APPLICANT: Spytek, Kimberly
: APPLICANT: Shenoy, Suresh
: APPLICANT: Li, Li
: APPLICANT: Gusev, Vladimir
: APPLICANT: Grosse, William
: APPLICANT: Alsobrook, John
: APPLICANT: Lepley, Denise
: APPLICANT: Burgess, Catherine
: APPLICANT: Gerlach, Valerie
: APPLICANT: Ellerman, Karen
: APPLICANT: MacDougall, John
: APPLICANT: Stone, David
: APPLICANT: Smithson, Glenda
: TITLE OF INVENTION: Novel Proteins and Nucleic Acids Encoding Same
: FILE REFERENCE: 21402-074 US
: CURRENT APPLICATION NUMBER: US/09/918,779
: CURRENT FILING DATE: 2001-07-30
: PRIOR APPLICATION NUMBER: 60/221,409
: PRIOR FILING DATE: 2000-07-28
: PRIOR APPLICATION NUMBER: 60/222,840
: PRIOR FILING DATE: 2000-08-04

```

	/	PRIOR APPLICATION NUMBER:	60/223,752	
	/	PRIOR FILING DATE:	2000-08-08	
	/	PRIOR APPLICATION NUMBER:	60/223,762	
	/	PRIOR FILING DATE:	2000-08-08	
	/	PRIOR APPLICATION NUMBER:	60/223,770	
	/	PRIOR FILING DATE:	2000-08-08	
	/	PRIOR APPLICATION NUMBER:	60/223,769	
	/	PRIOR FILING DATE:	2000-08-08	
	/	PRIOR APPLICATION NUMBER:	60/225,146	
	/	PRIOR FILING DATE:	2000-08-14	
	/	PRIOR APPLICATION NUMBER:	60/225,392	
	/	PRIOR FILING DATE:	2000-08-15	
	/	PRIOR APPLICATION NUMBER:	60/225,470	
	/	PRIOR FILING DATE:	2000-08-15	
	/	PRIOR APPLICATION NUMBER:	60/225,697	
	/	PRIOR FILING DATE:	2000-08-16	
	/	PRIOR APPLICATION NUMBER:	60/263,662	
	/	PRIOR FILING DATE:	2001-02-01	
	/	PRIOR APPLICATION NUMBER:	60/281,645	
	/	PRIOR FILING DATE:	2001-04-05	
	/	NUMBER OF SEQ ID NOS:	61	
	/	SOFTWARE:	PatentIn Ver. 2.1	
	/	SEQ ID NO 1		
	/	LENGTH:	2752	
	/	TYPE:	DNA	
	/	ORGANISM:	Homo sapiens	
	/	US-09-918-779-1		
		Query Match	92.9%; Score 2676.4; DB 13; Length 2752;	
		Best Local Similarity	98.9%; Pred. No. 0;	
		Matches 2728; Conservative	0; Mismatches 21; Indels 9; Gaps 3;	
QY	42	CCGCGGGCCCCCGCGGCCCGCCCGCCTTGC	CGCGCGCATAGTCCTCGCGCTTTGGCTCCGCGCTCGGGTGCC	161
Db	1	CGCGGGGCCCGCGCGCGGCCCGCCCGCTG	CCTCGCGCGCATAGTCCTCGCGCTTTGGCTCCGCGCTCGGGTGCC	120
QY	102	GGCCTGTGGCCAGCGCTCTCTGGGCATAGT	TCTTCAAGTGCAACGGGGAGTGGGTGCGCCAGGTG	341
Db	61	GGCCTGTGGCCAGCGCTCTTGGGCATAGT	TCTTCAAGTGCAACGGGGAGTGGGTGCGCCAGGTG	300
QY	162	CAGCAGAGTGCCACGCTGGCCCAACCAGT	GCTGGTGGTGCACCCGGAAGTCTGCTTCCCAC	221
Db	121	CAGCAGAGTGCCACGCTGGCCCAACCAGT	GCTGGTGGTGCACCCGGAAGTCTGCTTCCCAC	180
QY	222	TTCTGTGTGGAGCCCGAGGATGTGTACAT	CGTCAAGAAACAAGCCAGTGCCTTGTGTGC	281
Db	181	TTCTGTGTGGAGCCCGAGGATGTGTACAT	CGTCAAGAAACAAGCCAGTGCCTTGTGTGC	240
QY	282	AAGCCCTGCCC CGCACGAGATCTTCTT	CAAGTGCAACGGGGAGTGGGTGCGCCAGGTG	341
Db	241	AAGCCCTGCCC CGCACGAGATCTTCTT	CAAGTGCAACGGGGAGTGGGTGCGCCAGGTG	300
QY	342	GACCACGTGATCGAGCGCACACAGACGG	AGAGTGGTAGCCGACCATGAGAGTCCGC	401
Db	301	GACCACGTGATCGAGCGCACACAGACGG	AGAGTGGTAGCCGACCATGAGAGTCCGC	360
QY	402	ATTAAATGCTCAAAGGCAGCAGGTGCGA	AGAGTCTTCGGGCTGGAGAAACTGCTGCCAG	461
Db	361	ATTAAATGCTCAAAGGCAGCAGGTGCGA	AGAGTCTTCGGGCTGGAGAAACTGCTGCCAG	420
QY	462	TGCGTGCCATGGAGCTCTTCGGGCACCA	CCAAGAGTCAGAAGGCCTTACATCCGCATAGCC	521
Db	421	TGCGTGCCATGGAGCTCTTCGGGCACCA	CCAAGAGTCAGAAGGCCTTACATCCGCATAGCC	480
QY	522	AGATTGGCAAGAACCTTCAGCAGGAGCG	CTGGCCCAAGGAGGTGTCTTGAGCAGGCG	581
Db	481	AGATTGGCAAGAACCTTCAGCAGGAGCG	CTGGCCCAAGGAGGTGTCTTGAGCAGGCG	540
QY	582	ATCGTGTGCTCCCTCGCTCCACCGGAG	GCATCCCTCCAGCGAGGTGAGTGGTCCGG	641
Db	541	ATCGTGTGCTCCCTCGCTCCACCGGAG	GCATCCCTCCAGCGAGGTGAGTGGTCCGG	600
QY	642	AACGAGGACTGGTGGACCGCTCCCTTG	AGCCCAATGTATACATCAACGCGGACACAGC	701

1362 GACCTCAG--CACACACACACCTACAGGGCAGTCTGTGTCCCGCGCAGATGGGCC 1418
1318 GACCTCAGACACACACACACCTACAGGGCAGTCTGTGTCCCGCGCAGATGGGCC 1377
1419 AGCCCAAGTTCAGCTCACAATGGGACCTGTCTAGCCCTCGGTGGCGCGCCGAC 1478
1378 AGCCCAAGTTCAGCTCACAATGGGACCTGTCTAGCCCTCGGTGGCGCGCCGAC 1437
1479 ACACCTGACACACAGCTCTCCACCTCTGAGGCGGAGGATTCGTCTCCGCTCTCCACC 1538
1438 ACACCTGACACACAGCTCTCCACCTCTGAGGCGGAGGATTCGTCTCCGCTCTCCACC 1497
1539 CAGAACTACTTCGGCTCTCCGCTCTCCGCTCTGAGGCGGAGGATTCGTCTCCGCTCTCCACC 1598
1498 CAGAACTACTTCGGCTCTCCGCTCTCCGCTCTGAGGCGGAGGATTCGTCTCCGCTCTCCACC 1557
1599 TTCTCTCGGGCGCGCTGATGATCTCTAATACAGGTATCAGCTCTCTCATCCCGCCAGAT 1658
1558 TTCTCTCGGGCGCGCTGATGATCTCTAATACAGGTATCAGCTCTCTCATCCCGCCAGAT 1617
1659 GCCATACCCCGGAGGAGATCTATGAGATCTACCTCAGCTCACAAGCCGGAAGACGTG 1718
1618 GCCATACCCCGGAGGAGATCTATGAGATCTACCTCAGCTCACAAGCCGGAAGACGTG 1677
1719 AGTTGCGCCCTAGCTGGCTGTGAGACCTGTGAGTCCCATCGTTAGCTGTGACCCCT 1778
1678 AGTTGCGCCCTAGCTGGCTGTGAGACCTGTGAGTCCCATCGTTAGCTGTGACCCCT 1737
1779 GCGCTCTCTCACCAGGCGCAGTCTCTCTGATGAGACCTGTGAGGAGCCGACCCCT 1838
1738 GCGCTCTCTCACCAGGCGCAGTCTCTCTGATGAGACCTGTGAGGAGCCGACCCCT 1797
1839 GACAGCTGGAGCTCGCTCTCAAAAGCAGTCTGTGAGGCGAGCTGTGAGGAGGATGTG 1898
1798 GACAGCTGGAGCTCGCTCTCAAAAGCAGTCTGTGAGGCGAGCTGTG--AGGATGTG 1854
1899 GTGACCTGGGAGAGGCGCCCTCCAGCTCTACTACTGAGTGTGAGGAGGAGGAGGAGGAG 1958
1855 CTGACCTGGGAGAGGCGCCCTCCAGCTCTACTACTGAGTGTGAGGAGGAGGAGGAGGAG 1914
1959 TGTAGCTCTTACCCAGCAGCTGGGCGCTTTGCGCTGTGAGGAGGAGGAGGAGGAGGAG 2018
1915 TGTAGCTCTTACCCAGCAGCTGGGCGCTTTGCGCTGTGAGGAGGAGGAGGAGGAGGAG 1974
2019 GCTGCGCCCAAGCGCTCAAGCTGTCTTTGCGCGGCTGTGAGGAGGAGGAGGAGGAGGAG 2078
1975 GCTGCGCCCAAGCGCTCAAGCTGTCTTTGCGCGGCTGTGAGGAGGAGGAGGAGGAGGAG 2034
2079 TACCAATCGGCTTACTGCTGTGATGACACCCAGATGCTCAAGGAGGAGGAGGAGGAGGAG 2138
2035 TACCAATCGGCTTACTGCTGTGATGACACCCAGATGCTCAAGGAGGAGGAGGAGGAGGAG 2094
2139 CTGGAAGCAGCTGGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 2198
2095 CTGGAAGCAGCTGGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 2154
2199 AGTTACCAACCTGGCGCTTATCCATCCAGATGTGCGGAGCTCTCTGTGGAAGAGTAAG 2258
2155 AGTTACCAACCTGGCGCTTATCCATCCAGATGTGCGGAGCTCTCTGTGGAAGAGTAAG 2214
2259 CTCTTGTGAGCTACAGGAGATCCCTTTTATCATCTGGAATGCGACGAGCGGTAC 2318
2215 CTCTTGTGAGCTACAGGAGATCCCTTTTATCATCTGGAATGCGACGAGCGGTAC 2274
2319 TTGCACTGCACTTACCCCTGGAGGAGTGTGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 2378
2275 TTGCACTGCACTTACCCCTGGAGGAGTGTGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 2334
2379 CTGTGGGTGTGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 2438
2335 CTGTGGGTGTGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 2394

2439 AAGGACACAAAGTTTGTGAGCTGTGCTCTGAGAGTGAAGCGGGGTCCACGCCCTG 2498
2395 AAGGACACAAAGTTTGTGAGCTGTGCTCTGAGAGTGAAGCGGGGTCCACGCCCTG 2454
2499 GTGGGCCCAAGTGTCTTCAAGATCCCTTCTCATTTGGCAGAAATATTTCCAGCCTG 2558
2455 GTGGGCCCAAGTGTCTTCAAGATCCCTTCTCATTTGGCAGAAATATTTCCAGCCTG 2514
2559 GACCCACCTGTAGCGGGGTGCGACTGGGGGACTCTGGCCAGAAACTCCACCTGGAC 2618
2515 GACCCACCTGTAGCGGGGTGCGACTGGGGGACTCTGGCCAGAAACTCCACCTGGAC 2574
2619 AGCCATCTCAGCTTCTTTGCTCTCAAGCCAGCCAGCCAGCCAGCCAGCCAGCCAG 2678
2575 AGCCATCTCAGCTTCTTTGCTCTCAAGCCAGCCAGCCAGCCAGCCAGCCAGCCAG 2634
2679 GAGGCGGCGACTTCCCAACCGGAACTCAGCAGCTGGCTGACAGCTGGCTGGACTG 2738
2635 GAGGCGGCGACTTCCCAACCGGAACTCAGCAGCTGGCTGACAGCTGGCTGGACTG 2694
2739 GGCACGCGAGCGCTGGCTCTTTCAGAGTGTGCGAGGCTGAGTGTGAGGCGCGGAC 2796
2695 GGCACGCGAGCGCTGGCTCTTTCAGAGTGTGCGAGGCTGAGTGTGAGGCGCGGAC 2752

RESULT 4

US-10-311-623-13
; Sequence 13, Application US/10311623
; Publication No. US20040023244A1
; GENERAL INFORMATION:
; APPLICANT: INCYTE GENOMICS, INC.; GRIFFIN, Jennifer A.
; APPLICANT: KALLICK, Deborah A.; TRIBOULEY, Catherine M.
; APPLICANT: YUE, Henry; NGUYEN, Daniel B.
; APPLICANT: TANG, Y. Tom; LAL, Preeti G.
; APPLICANT: POLICKY, Jennifer L.; AZIMZAI, Yalda
; APPLICANT: LU, Dyoung Aina M.; GRAUL, Richard C.
; APPLICANT: YAO, Monique G.; BURFORD, Neil
; APPLICANT: HAFALIA, April J. A.; BAUGHN, Mariah R.
; APPLICANT: BANDMAN, Olga; ARVIZU, Chandra S.
; APPLICANT: YANG, Junming; XU, Yuming
; APPLICANT: GANDHI, Ameena R.; WARREN, Bridget A.
; APPLICANT: DING, Li; SANJANWALA, Madhusudan M.
; APPLICANT: DUGGAN, Brendan M.; LU, Yan
; TITLE OF INVENTION: RECEPTORS
; FILE REFERENCE: PF-0793 USN
; CURRENT APPLICATION NUMBER: US/10/311,623
; PRIOR FILING DATE: 2002-12-17
; PRIOR APPLICATION NUMBER: US 01/19942
; PRIOR FILING DATE: 2001-06-21
; PRIOR APPLICATION NUMBER: US 60/214,027
; PRIOR FILING DATE: 2000-06-21
; PRIOR APPLICATION NUMBER: US 60/228,045
; PRIOR FILING DATE: 2000-08-25
; PRIOR APPLICATION NUMBER: US 60/255,104
; PRIOR FILING DATE: 2000-12-12
; NUMBER OF SEQ ID NOS: 24
; SOFTWARE: PERL Program
; SEQ ID NO 13
; LENGTH: 3580
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: misc feature
; OTHER INFORMATION: Incyte ID No. US20040023244A1 6052371CB1
US-10-311-623-13

Query Match 83.4%; Score 2402.6; DB 17; Length 3580;
Best Local Similarity 93.4%; Pred. No. 0;
Matches 2615; Conservative 0; Mismatches 9; Indels 177; Gaps 3;

QY 84 GCCATGCGCTCGCGCCCGCCCTGTGCGCAGCGCTCTCTGCGCATAGTCTCGCGCTGG 143
Db 1 GCCATGCGCTCGCGCCCGCCCTGTGCGCAGCGCTCTCTGCGCATAGTCTCGCGCTGG 60

QY	144	CTCCGCGGCTCGGGTGCCTCCAGCAGAGTGCACCGTGGCCAAACCCAGTGCCTGGTGCCAC	203
DB	61	CTCCGCGGCTCGGGTGCCTCCAGCAGAGTGCACCGTGGCCAAACCCAGTGCCTGGTGCCAC	120
QY	204	CCGACCTGCTTCCCACTTCTCTGGTGGAGCCGAGATGTGTACATCTCTCAAGAACAG	263
DB	121	CCGACCTGCTTCCCACTTCTCTGGTGGAGCCGAGATGTGTACATCTCTCAAGAACAG	180
QY	264	CCAGTGCCTGTGTGTGCAAGCCGTGCCCGCAACGCAGATCTTCTTAAGTGCACCGG	323
DB	181	CCAGTGCCTGTGTGTGCAAGCCGTGCCCGCAACGCAGATCTTCTTAAGTGCACCGG	240
QY	324	GAGTGGGTGCGCAGGTGGACACAGTATCGAGCGCAGACAGACGGAGCAGTGGTGAG	383
DB	241	GAGTGGGTGCGCAGGTGGACACAGTATCGAGCGCAGACAGACGGAGCAGTGGTGAG	300
QY	394	CCGACCATGGAGGTCCGCAATTAATCTCAAGGCAGCAGTCCGAGAAGGTGTTCGGGCTG	443
DB	301	CCGACCATGGAGGTCCGCAATTAATCTCAAGGCAGCAGTCCGAGAAGGTGTTCGGGCTG	360
QY	444	GAGGAATACTGGTCCAGTGCCTGGCATCGAGTCTCTCGGGACACCAAGAGTCCAGAAG	503
DB	361	GAGGAATACTGGTCCAGTGCCTGGCATCGAGTCTCTCGGGACACCAAGAGTCCAGAAG	420
QY	504	GCCTACATCCGCATAGCCAGATTGCGCAAGAACTTTCGAGCAGAGCCGTGCCCAAGAG	563
DB	421	GCCTACATCCGCATAGCCATTTTGGCAGAACTTTCGAGCAGAGCCGTGCCCAAGAG	480
QY	564	GTGTCCCTGGAGCAGGGCATCGTGTGCCTGCCGTCCACCGAGGGCATCCCTCCAGCC	623
DB	481	GTGTCCCTGGAGCAGGGCATCGTGTGCCTGCCGTCCACCGAGGGCATCCCTCCAGCC	540
QY	624	GAGGTGGAGTGCCTCCGAAACGAGACCTGTGTGACCCGTCTCCCTGGACCCCAATGTATAC	683
DB	541	GAGGTGGAGTGCCTCCGAAACGAGACCTGTGTGACCCGTCTCCCTGGACCCCAATGTATAC	600
QY	684	ATCACGGGGAGCAGACGCTGTGTGTGCGACAGGCCCGCTTGTGTGACACGGCCAACTAC	743
DB	601	ATCACGGGGAGCAGACGCTGTGTGTGCGACAGGCCCGCTTGTGTGACACGGCCAACTAC	660
QY	744	ACCTGGTGGCCAAAGAACATCGTGGCAGCTCCCGCAGCGCTCCGTGCTGTCTATCGTC	803
DB	661	ACCTGGTGGCCAAAGAACATCGTGGCAGCTCCCGCAGCGCTCCGTGCTGTCTATCGTC	720
QY	804	TACGTGAACCGTGGTGTGTGACAGTGGACCGAGTGGTCCGTCTGCGAGGCCAGCTGGGG	863
DB	721	TAC-----TAC-----	723
QY	864	CGCGCTGGCAAGAACGGAGCGGAGGTGCACCAACCGCGCCCTCTCAACGGGGCGCT	923
DB	724	-----	723
QY	924	TTCTGTGAGGGCAGAAATGTCCATGACCGCACCGTCTCCTCTCTGTGTGTCTGTGTGAC	983
DB	724	-----GTGGAC-----	729
QY	984	GGACGTGGAGCCGTGGAGCAAGTGTGCGCCTGTGGCTGGAGTGCACCCACTGGCGG	1043
DB	730	GGACGTGGAGCCCGTGGAGCAAGTGTGCGCCTGTGGCTGGAGTGCACCCACTGGCGG	789
QY	1044	AGCCGTGAGTGTCTGTACCCAGCACCCCGCAACGGAGGGGAGGAGTGCACGGGCACTGAC	1103
DB	790	AGCCGTGAGTGTCTGTACCCAGCACCCCGCAACGGAGGGGAGGAGTGCACGGGCACTGAC	849
QY	1104	CTGACACCCGCAACTGTACAGTGAACCTCTGTGTACACAGTGTCTTGTGGCCTTAGGAC	1163
DB	850	CTGACACCCGCAACTGTACAGTGAACCTCTGTGTACACAGTGTCTTGTGGCCTTAGGAC	909
QY	1164	GTGGCCCTCTATGTGGGCCTCATCGCGTGGCCGTCTGCTGTGCTGCTGCTGCTGCTGTC	1223
DB	910	GTGGCCCTCTATGTGGGCCTCATCGCGTGGCCGTCTGCTGTGCTGCTGCTGCTGCTGTC	969

QY	1224	CTCATCTCTCGTTTATTGCGCGAAGAAGAGGGGTGGACTCAGATGTGGCTGACTCGTCC	1281
Db	970	CTCATCTCTCGTTTATTGCGCGAAGAAGAGGGGTGGACTCAGATGTGGCTGACTCGTCC	1029
QY	1284	ATTCTCACCTCAGGCTTCCAGCCCGTTCAGATCAAGCCCAAGCAGACAGAACCCCAT	1343
Db	1030	ATTCTCACCTCAGGCTTCCAGCCCGTTCAGATCAAGCCCAAGCAGACAGAACCCCAT	1089
QY	1344	CTGCTCACCATCCAGCCGGAGCTCAG---CACCAACCACCACTACCAGGGCAGTCTCTGT	1400
Db	1090	CTGCTCACCATCCAGCCGGAGCTCAGCACACACACCACTTACAGGCGAGTCTCTGT	1149
QY	1401	CCCCGGCAGGATGGGCCCCAGCCCCAAGTTCAGTCTCAAAATGGGCACCTGCTCAGCCCC	1460
Db	1150	CCCCGGCAGGATGGGCCCCAGCCCCAAGTTCAGTCTCAAAATGGGCACCTGCTCAGCCCC	1209
QY	1461	CTGGGTGGGGCGGCCACACACTGCACACAGCTTCTCCACCTCTGAGGCGGAGGAGTTC	1520
Db	1210	CTGGGTGGGGCGGCCACACACTGCACACAGCTTCTCCACCTCTGAGGCGGAGGAGTTC	1269
QY	1521	GTCTCCCGCCTCTCCACCAGAACTACTTCGGCTCCCTGCCCGAGSCACCAACAATG	1580
Db	1270	GTCTCCCGCCTCTCCACCAGAACTACTTCGGCTCCCTGCCCGAGSCACCAACAATG	1329
QY	1581	ACCTATGGGACCTTCAACTTCTCGGGGCGGCTGATGATCCCTAATACAGGTATCAGC	1640
Db	1330	ACCTATGGGACCTTCAACTTCTCGGGGCGGCTGATGATCCCTAATACAGGAATCAGC	1389
QY	1641	CTCCTCATCCCCCAGATGCCATACCCCGAGGGAAGATCTATGAGATCTACTTCACGCTG	1700
Db	1390	CTCCTCATCCCCCAGATGCCATACCCCGAGGGAAGATCTATGAGATCTACTTCACGCTG	1449
QY	1701	CACAAGCCGGAAGAGCTGAGGTTGCCCTAGCTTGGCTGTCCAGACCTGCTGAGTCCCATC	1760
Db	1450	CACAAGCCGGAAGAGCTGAGGTTGCCCTAGCTTGGCTGTCCAGACCTGCTGAGTCCCATC	1509
QY	1761	GTTAGCTGTGAACCCCTGGCGTCTGCTCACCCGGCCAGTCACTCTGGCTATGGACCAC	1820
Db	1510	GTTAGCTGTGAACCCCTGGCGTCTGCTCACCCGGCCAGTCACTCTGGCTATGGACCAC	1569
QY	1821	TGTGGGAGCCAGCCCTGACAGTGGAGCCTGGCCTCAAAAGCAGTGTGCGAGGGC	1880
Db	1570	TGTGGGAGCCAGCCCTGACAGTGGAGCCTGGCCTCAAAAGCAGTGTGCGAGGGC	1629
QY	1881	AGCTGGCAGCAGGATGTGCTGACCTGGGCGAGAGGCGCCCTCCACCTCTACTACTGC	1940
Db	1630	AGCTGGG---AGGATGTGCTGCATCTGGGCGAGAGGCGCCCTCCACCTCTACTACTGC	1686
QY	1941	CAGCTGGAGCGCAGTGCTCTACGTCTTCCCGAGCAGCTGGGCGGCTTTGCCCTGGTG	2000
Db	1687	CAGCTGGAGCGCAGTGCTCTACGTCTTCCCGAGCAGCTGGGCGGCTTTGCCCTGGTG	1746
QY	2001	GGAGAGCCCTCAGCTGGCTGGCGGCAAGCGCCTCAAGCTGTTCTGTTTGGCGCGGTG	2060
Db	1747	GGAGAGCCCTCAGCGTGGCTGGCGGCAAGCGCCTCAAGCTGTTCTGTTTGGCGCGGTG	1806
QY	2061	GCCTGCACTCCCTCGAGTACAACTCCGGTCTACTGCTCGATGACACCAACATGCA	2120
Db	1807	GCCTGCACTCCCTCGAGTACAACTCCGGTCTACTGCTCGATGACACCAACATGCA	1866
QY	2121	CTCAAGAGGTGGTGCAGCTGGAGAGCAGCTGGGGGACAGCTCATCCAGGAGCCAGG	2180
Db	1867	CTCAAGAGGTGGTGCAGCTGGAGAGCAGCTGGGGGACAGCTCATCCAGGAGCCAGG	1926
QY	2181	GTCTTGCACTTCAAGAGCAGTTACCAACAACCTGGCGCTTATCCATCCAGATGTGCCAGC	2240
Db	1927	GTCTTGCACTTCAAGAGCAGTTACCAACAACCTGGCGCTATCCATCCAGATGTGCCAGC	1986
QY	2241	TCCTGTGGAAGAGTAAGTCCCTGTACGTACCAAGGAGATCCCTTTTATACATCTGG	2300
Db	1987	TCCTGTGGAAGAGTAAGTCCCTGTACGTACCAAGGAGATCCCTTTTATACATCTGG	2046
QY	2301	AATGGCACGCGAGCGGTACTTTCAGCTGCACCTTCAACCTGGAGCGGTGTGAGGCCAGCACT	2360

Db 2047 AATGACGCGAGCGTACTTGACATGACCTTACCTTGAGCGTGTAGCCCGAGCACT 2106
Qy 2361 AGTGACCTGGCGCTGCAAGCTGTGGGTGTGGCAGGTGGAGGGCGACGGGCGAGCTTCAGC 2420
Db 2107 AGTGACCTGGCGCTGCAAGCTGTGGGTGTGGCAGGTGGAGGGCGACGGGCGAGCTTCAGC 2166
Qy 2421 ATCACTTCAACATCACCAAGACACACAGGTTTGTGAGCTGTGAGCTGTGAGTGTGAGAGTGAA 2480
Db 2167 ATCACTTCAACATCACCAAGACACACAGGTTTGTGAGCTGTGAGCTGTGAGTGTGAGAGTGAA 2226
Qy 2481 GCGGGGGTCCAGCGCTGTGGGCGCCAGTGTCTTCAAGATCCCTTCTCAATCGGCAG 2540
Db 2227 GCGGGGGTCCAGCGCTGTGGGCGCCAGTGTCTTCAAGATCCCTTCTCAATCGGCAG 2286
Qy 2541 AAGATAATTTCCAGCGCTGACCCACCTGTAGCGGGGTGCGGACTTGGCGGACTTGGCC 2600
Db 2287 AAGATAATTTCCAGCGCTGACCCACCTGTAGCGGGGTGCGGACTTGGCGGACTTGGCC 2346
Qy 2601 CAGAACTCCAGCTGGACAGCATCTCAGCTTCTTGGCTTCCAGCCAGCCCGCCAGCC 2660
Db 2347 CAGAACTCCAGCTGGACAGCATCTCAGCTTCTTGGCTTCCAGCCAGCCCGCCAGCC 2406
Qy 2661 ATGATCTCAACTGTGGAGGGCGGCACTTCCCAACCGCAACCTCAGCCAGCTGGCT 2720
Db 2407 ATGATCTCAACTGTGGAGGGCGGCACTTCCCAACCGCAACCTCAGCCAGCTGGCT 2466
Qy 2721 GCAGCAGTGGCTGGAGTGGCCAGCCAGCAGCTGTGGCTTTTCAAGTGTGGAGGCTGAG 2780
Db 2467 GCAGCAGTGGCTGGAGTGGCCAGCCAGCAGCTGTGGCTTTTCAAGTGTGGAGGCTGAG 2526
Qy 2781 TGCTGAGCGCGCCAGCGCCGACCTACACTCTCAGCAGCTTGGCAGCCCGCCAGGAG 2840
Db 2527 TGCTGAGCGCGCCAGCGCCGACCTACACTCTCAGCAGCTTGGCAGCCCGCCAGGAG 2586
Qy 2841 AGCAGAGCGCGACAGGGCGCTTCCCAACCGGGGAGA 2881
Db 2587 AGCAGAGCGCGACAGGGCGCTTCCCAACCGGGGAGA 2627

RESULT 5
US-09-933-261-1
; Sequence 1, Application US/09933261
; Publication No. US20030040046A1
; GENERAL INFORMATION:
; APPLICANT: Tessier-Lavigne, Marc
; Leonardo, E. David
; Hink, Lindsey
; Masu, Masayuki
; Kazuko, Keino-Masu
; TITLE OF INVENTION: Netrin Receptors
; NUMBER OF SEQUENCES: 8
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: SCIENCE & TECHNOLOGY LAW GROUP
; STREET: 268 BUSH STREET, SUITE 3200
; CITY: SAN FRANCISCO
; STATE: CALIFORNIA
; COUNTRY: USA
; ZIP: 94104
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent In Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/933,261
; FILING DATE: 20-Aug-2001
; CLASSIFICATION: <unknown>
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 08/808,982
; FILING DATE: <unknown>
; ATTORNEY/AGENT INFORMATION:
; NAME: OSMAN, RICHARD A

; REGISTRATION NUMBER: 36,627
; REFERENCE/DOCKET NUMBER: UC96-217
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (415) 343-4341
; TELEFAX: (415) 343-4342
; INFORMATION FOR SEQ ID NO: 1:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 3014 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: double
; TOPOLOGY: linear
; MOLECULE TYPE: cDNA
; SEQUENCE DESCRIPTION: SEQ ID NO: 1:
US-09-933-261-1

Query Match 77.6%; Score 2235.8; DB 10; Length 3014;
Best Local Similarity 88.5%; Pred. No. 0;
Matches 2474; Conservative 0; Mismatches 312; Indels 11; Gaps 4;

Qy 87 ATGGCGCTCGGCGCGCGCTGTGGCAGCGCTCTCTGGGCATAGTCTCTCGCGCTTGGCTC 146
Db 1 ATGGCGCTCGGCGCGCGCTGTGGCAGTGTCTCTGGGCATAGTCTCTCGCGCTTGGCTC 60

Qy 147 CCGCGCTCGGCTCGGCGCGCGCTGTGGCAGAGTGCACCGTGGCCCAACCCAGTGTCTGTGCCAACCCG 206
Db 61 CGTGGTTCGGGTGCCGAGAGTGCACCGTGGCCCAATCCAGTGTCCCGTGTCCCAACCCG 120

Qy 207 GACCTGCTTCCCACTTCTCTGTGGCAGCGCGCGAGGATGTGTACATCGTCAAGAACAGGCCA 266
Db 121 GACCTGCTTCCCACTTCTCTGTGGCAGCGCTGTGGCAGCGTGTATTCATTTCAAGAACAGGCCG 180

Qy 267 GTGCTGCTTGTGTGCAAGCGCGTGGCGCGCAGCGAGATCTTCTTCAAGTGCACACGGGAG 326
Db 181 GTGTTGTTGTGTGCAAGCGCTGTGGCTGCCACCGAGATCTTCTTCAAGTGCATGGGAA 240

Qy 327 TGGGTGCGCAAGTGTGACCACTGTGATCGAGCGCGAGCACAGCGGAGCAGTGTGTAGCCG 386
Db 241 TGGGTGCGCAGTGTGATCACGTAATTGAACGCGAGCACCGCAGCAGCAGCGGATTTGCCA 300

Qy 387 ACCATGAGGTCCGCAATTAATGTCTCAAGCGAGCAGTGTGAGAGGTGTTCGGGCTGGAG 446
Db 301 ACCATGAGGTCCGTAATCAACGTATCGAGCGAGCAGGTAGAGAAAGTGTTCGGGCTGGAG 360

Qy 447 GAATACTGTGCGCAGTGTGCGCATGGAGCTCTCTCGGCGCACCCAGAGTGTAGAGGCC 506
Db 361 GAATACTGTGCGCAGTGTGCGCATGGAGCTCTCTCGGCGTACACCAAAAGTGTAGAGGCC 420

Qy 507 TACATCCGATAGCCAGATTCGCAAGAACTTCGAGCAGGAGCGCTGTGCCAAGGAGGTG 566
Db 421 TACATCCGATAGCCATTTTTCGCAAGAACTTTGAGCAGGAGCCACTGTGCCAAGGAGGTG 480

Qy 567 TCCCTGAGCGAGCGCATCGTGTCCCTGCGCTCCACCGGAGGCGATCCCTCCAGCGCGAG 626
Db 481 TCATCGAGCAGGCGCATTTACTTACCTTGTGCGCGCGCGCGAGAGGATTCCTCCAGCTGAG 540

Qy 627 GTGAGTGTGCTCCGGAACGAGGACCTGTGGAGACCGCTTCCTTGGACCGCCCAATATATAC 686
Db 541 GTGAGTGTGCTTCGAATGAGGACCTGTGGAGACCGCTTCCTTGGATCCCAATGTATATC 600

Qy 687 ACCTGAGCAGCAGCGCTGTGGCAGCGCGCGCTTCCTGACACGGCGCACTACAC 746
Db 601 ACCTGAGCAGCAGCGCTGTGGCAGCGCGCGCTTCCTGACACGGCGCACTACAC 660

Qy 747 TCGTGTGCGCAAGAACATTCGTGGCAGCGCGCGCTTCCTGCTGTGTATCTGTCTAC 806
Db 661 TGTGTGCGCAAGAACATTCGTAGCGCGCGCGAGGACCTCTGAGCGGCTCATTTTAT 720

Qy 807 GTGAACCGTGTGCTGCGAGCGAGTGTGTCCGTGTGACGCGCAGCTGTGGCGCG 866
Db 721 GTGAACCGTGTGCTGCGAGCGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 780

Qy 867 GGTGTGAGAAACGAGCGCGAGTGTGACCAACCGCGCGCTCTCAACGGGGCGCTTTC 926

QY 1644 CTCATCCCCCAGATGCCATACCCCGAGGGAAGATCTATGAGATCTACCTCAGCTGCAC 1703
Db 1558 CTCATACCCCGGATGCCATCCCGGAGGAAGATCTACAGATCTACCTCAGCTGCAC 1617
QY 1704 AAGCCGGAAGAGTGTAGGTTGCCCTTAGCTGGTGTGAGACCTCTGAGTCCCATCGTT 1763
Db 1618 AAGCCGAAGAAGTGTAGGTTGCCCTTAGCTGGTGTGAGACCTCTGAGTCCCATCGTT 1677
QY 1764 AGCTGTGGACCCCTCGGCTCTGCTCACCAGCCAGTCTATCTGCTATGACCACTGT 1823
Db 1678 AGCTGTGGCCCGCCAGAGTCTCTGCTACCCGCGCAGTCTATCTGCAATGGAACCACTGT 1737
QY 1824 GGGAGCCAGCCCTGACAGCTGGAGCTCGGCTCAAAAAGCAGTCGTCGAGGCGAGC 1883
Db 1738 GGAGAGCCAGCCCTGACAGCTGGAGTCTGGGCTCAAAAAGCAGTCTCTGAGGCGAGT 1797
QY 1884 TGGGAGCAGGATGTGTGACCTGGCGGAGGAGGCGCTTCCACCTCTACTACTGCCAG 1943
Db 1798 TGGG---AGGATGTGTGACCTTGTGAGGAGTCACTTCCACCTCTACTACTGCCAG 1854
QY 1944 CTGAGGCGCAGTGTCTGCTACGTCCTTACCGAGCAGCTGGCGGCTTGGCCTGTGGA 2003
Db 1855 CTGAGGCGCGGCGCTGTCTGCTTTCAGGAGCAGCTGGGCGGCTTGGCCTGTGGA 1914
QY 2004 GAGGCCCTCAGCTGTGCTGCCCAAGCGCTCAAGCTCTCTGTTTGGCGGCTGGCC 2063
Db 1915 GAGGCCCTCAGCTGTGCTGCCCAAGCGCTCAGGCTCTCTGTTTGGCTTCCGTTGCC 1974
QY 2064 TGCACCTCCCTCGAGTACAACTCGGCTCTACTGCTGCTGATGACACCCAGTGCATC 2123
Db 1975 TGTAGCTCCCTTGAATACAACTCGGCTCTACTGCTTACAGACACCCAGCAGCTCTC 2034
QY 2124 AAGGAGGTGGTGCAGCTGGAGAGCAGCTGGGCGGAGCAGTCTATCAGAGGACCGGTC 2183
Db 2035 AAGGAGGTGGTGCAGCTGGAGAGCAGCTAGGTGGACAGCTGATCCAGGAGCCTCGGTC 2094
QY 2184 CTGCACTTCAAGGACAGTTTACCAACCTTGCCTTATCCATCCAGATGTGCCAGCTCC 2243
Db 2095 CTGCACTTCAAGGACAGTTTACCAACCTTGCCTTATCCATCCAGAGCTGCCAGCTCC 2154
QY 2244 CTGTGAAGAGTAAGTCTTGTGAGTACAGGAGATCCCTTTTATCACTCTGGAAT 2303
Db 2155 CTGTGAAGAGTAAGTCTTGTGAGTACAGGAGATCCCTTTTATCACTCTGGAAT 2214
QY 2304 GGCAGCGAGCTACTTGCAGTGCACCTTCAACCTGGAGCGTGTGAGCCAGCAGTACT 2363
Db 2215 GGCAGCGAGCTACTTGCAGTGCACCTTCAACCTGGAGCGGATCAACGCGAGCAGCAGC 2274
QY 2364 GACCTGGCTTCAAGCTGTGGTGTGGCAGGTGGAGGCGAGCGGAGAGCTTCAAGATC 2423
Db 2275 GACCTGGCTTCAAGCTGTGGTGTGGCAGGTGGAGGAGATGGCAGAGCTTCAAGATC 2334
QY 2424 AACTTCACATCACCAAGACACAGGTTTGTGAGCTGTGGTCTGGAGAGTGAAGCG 2483
Db 2335 AACTTCACATCACCAAGACACAGGTTTGTGAAATTTGTGGCTCTGGAGAGTGAAGCG 2394
QY 2484 GGGGTCCAGCCCTGTGGGCGCCAGTGCCTTCAAGATCCCTTCTCAATTCGGCAGAG 2543
Db 2395 GGGGTCCAGCCCTGTGGGCGCCAGTGCCTTCAAGATCCCTTCTCAATTCGGCAGAG 2454
QY 2544 ATAATTTCCAGCTTGGACCCACCTTGTAGGGGGGTGCGGAGTCTGGCGGAGTCTGGCCAG 2603
Db 2455 ATCATGGCAGTCTGGACCCACCTTGCAGCGGGGCGCGAGTGGAGAACTCTAGCCAG 2514
QY 2604 AAATCCACCTGGAGAGCAGTCTAGCTTTTGTCTTCAAGCCAGCCAGCCAGCCATG 2663
Db 2515 AAATCCACCTGGAGAGCAGTCTAGCTTTTGTCTTCAAGCCAGCCAGCCAGCCATG 2574
QY 2664 ATCTCAACCTGTGGAGCGCGGACCTTCCCAACGCGACCTTCAAGCAGCTGGCTGCA 2723
Db 2575 ATCTCAACCTGTGGAGCGCGGACCTTCCCAACGCGACCTTCCCAACGCGAGCTGGCAGCA 2634

QY 2724 GCAGTGGCTGGATGGCGCAGCCAGAGCTTGGCTCTTTCACAGTGTGGAGGCTGAGTGC 2783
Db 2635 GCTGTGGCGGAGTGGGCGCAACAGATGTGGCTCTTTCACGCTGTGGAGGCGGAGTGT 2694
QY 2784 TGAGGCGGCGCAGGCGCG--ACACCTTACACTCTCACCAGCTTTGGCACCCACCAAGACA 2841
Db 2695 TGAGACCAAGCAGGCGCGGTAATGCTTACCTTCTCACCAGCTTTGACACCTGCCAGGACA 2754
QY 2842 GGCAGAAAGCCGAGCAGGCGGCTTCCCAACACCGGGG 2878
Db 2755 GGCAAAACAGACAGGCGGCTTCCCAACACCGGGG 2791

RESULT 7

US-10-240-154-15
; Sequence 15, Application US/10240154
; Publication No. US20030175741A1
; GENERAL INFORMATION:
; APPLICANT: Cochran et al.
; FILE REFERENCE: SCHIZOPHRENIA RELATED GENES
; CURRENT APPLICATION NUMBER: US/10/240,154
; CURRENT FILING DATE: 2001-04-02
; PRIOR APPLICATION NUMBER: PCT/GB01/01486
; PRIOR FILING DATE: 2001-04-02
; NUMBER OF SEQ ID NOS: 34
; SOFTWARE: PatentIn version 3.2
; SEQ ID NO 15
; LENGTH: 2697
; TYPE: DNA
; ORGANISM: Rattus sp.
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (1)..(2697)
US-10-240-154-15

Query Match

Best Local Similarity 76.0%; Score 2189.4; DB 15; Length 2697;
Matches 2403; Conservative 0; Mismatches 291; Indels 9; Gaps 3;

QY 87 ATGGCCGTCCGCGCCGCGCTGTGGCAGCGCTCTGGGCATAGTCTCTGGCCTTGGCTC 146
Db 1 ATGGCCGTCCGCGCCGCGCTGTGGCAGTGTCTCTGGGCATAGTCTCTGGCCTTGGCTC 60
QY 147 CGGCTCGGTGGCCAGCAGAGTSCCACCCTGGCCAAACCCAGTGCCTGGTGCACACCG 206
Db 61 CGTGGTTCCGCTGGCCAGCAGAGTSCCACCCTGGCCAAATCCAGTGCCTGGTGCACACCG 120
QY 207 GACCTGTCTCCCACTTCTCTGGTGGAGCCCGAGGATGTGTACATCGTCAAGAACGCA 266
Db 121 GACCTGTCTCCCACTTCTCTGGTGGAGCCCGAGGATGTGTACATCGTCAAGAACGCG 180
QY 267 GTGCTGTGTGTGCAAGGCGCTGCGCCCGCAGCAGATCTTCTTCAAGTGAACGGGAG 326
Db 181 GTGTTGTGTGTGCAAGGCTGTGCTGCGCCAGCAGATCTTCTTCAAGTGAACGGGAG 240
QY 327 TGGTGGCCAGGTGGACCAAGTGTATGAGCGCAGCAGACGAGGAGCAGTGTGGAGCG 386
Db 241 TGGTGGCCAGGTGCATACGTAATTAAGCGCAGCAGCAGCAGCAGCAGCAGCAGCAG 300
QY 387 ACCATGAGGTCCCATTAATGTCTCAAGCAGCAGGTTCGAGAGGTGTTCGGCTGGAG 446
Db 301 ACCATGAGGTCCCATTAATGTCTCAAGCAGCAGGTTCGAGAGGTGTTCGGCTGGAG 360
QY 447 GAATACTGTGTCAGTGTGGCATGTGGCTCTCGGCGCACCAACAGAGTCAAGAGGCC 506
Db 361 GAATACTGTGTCAGTGTGGCATGTGGCTCTCGGCGTACCAACAGAGTCAAGAGGCC 420
QY 507 TACATCCGATAGCAGATTGCGCAAGAACTTCGAGCAGGAGCGCTGGCCAGAGGTG 566
Db 421 TACATCCGATAGCAGATTGCGCAAGAACTTCGAGCAGGAGCGCTGGCCAGAGGTG 480
QY 567 TCCCTGGAGCAGGCGCATCTGCTGCCCTTGGCTCCACCGGAGGCGCATCCCTTCAGCGGAG 626

	D _b	481	TCACTGGAGCAAGGCATTTGTAACCTTGTTCGCCCCCGAAGAAGAAATCCCCCAGTGAG	540
	Q _y	627	GTCGAGTGGCTCCGGAAACGAGGACTGTGTGGACCCTGTCCCTGGACCCGAATGTATACATC	686
	D _b	541	GTCGAGTGGCTTCGAAATGAGGACCTCGTGGACCCCTCCCTCGATCCCAATGTGTACATC	600
	Q _y	687	ACGCGGAGCACAGCCTGTGTGGCGACAGGCCCGCTTGTGACACGCGCAACTACACC	746
	D _b	601	ACGCGGAGCACAGCCTGTGTGGCGTAGTGTGTGAGCCCGCTGGCGACAGCGCCAATACACC	660
	Q _y	747	TGCGTGCCCAAGAACATCATGTGGCACGTGCGCGAGCGCCTCGCTGCTGTCAFGCTATAC	806
	D _b	661	TGTTGSCCAAGAACATCGTAGCCGTGCGGAAGCACCTCTCAGCGGTCAATTGTTTTAT	720
	Q _y	807	GTGMAACGGTGGGTGTCGACGTGGACCGAGTGTTCGTCGTGACGCCACAGCTGTGGCGC	866
	D _b	721	GTGAACGGTGGGTGTCGACGTGAGCTGAGTGTTCGTCGTGACGCCACAGCTGTGGCGCT	780
	Q _y	867	GGCTGGCAGAAAACGAGCCGAGCTGCACAAACCCGGCGCCTCTCAAACGGGGCGCTTTC	926
	D _b	781	GGCTGGCAGAAAACGAGCCGAGCTGCACAAACCCGGCACCTCTCAAACGGGGCGCTTTC	840
	Q _y	927	TGTGAGGGCAGAAATGTCATGACCGCACCGTCTCTCTCTGTCTGTCTGTGTGGACGGC	986
	D _b	841	TGTGAGGGCAGAAATGTCAGAAAAACGAGCTGCGCCACTCTG--TGCCAGTGGATGGG	897
	Q _y	987	AGCTGAGCCGCTGGAGCAAGTGTGGCTGTGGCTGTGACCTGCACCCCACCTGGCCGAGC	1046
	D _b	898	AGCTGGAGTTTCGTGGAGTAAGTGTGAGCTGTGGCTTGACTGCACCCACTGCGCGAGC	957
	Q _y	1047	CGTCAGTGTCTTGACCCAGCACCCCGCAACCGGAGGGAGGAGTGCCAGGSCACTGACCTG	1106
	D _b	958	CGCAGTGTCTTGACCCAGCACNCCCGCAATGGAGGTGAGGAGTGTGCGGTGTGTGACCTG	1017
	Q _y	1107	GACACCCGCAACTGTACAGTGACTCTGTGTACACAGTCTTCTTGCGCCCTGAGGACGTG	1166
	D _b	1018	GACACCCGCAACTGTACAGTGACTCTGTGCTGACACCGCTTCTTGCCGCCGAGGACGTG	1077
	Q _y	1167	GCCCTGTATGTGGGCTCATGCGCGTGGCGGTGTGCCCTGGTCTGTGCTGTGTGTGCTC	1226
	D _b	1078	GCTCTCTACATCGGCTTGTGCGCTGTGGCTTCTTCTGTGTGTGTGCTGTGGGCTT	1137
	Q _y	1227	ATCTCGTTTATTCGGGAACAAGAGSGGCTGGACCTCAGATGTGGCTGACTCGTCCATT	1286
	D _b	1138	GGACTCATTTACTGTGCAAGAAGAGGGCTGTGACTCCGATGTGGCCGACTGTGCAATC	1197
	Q _y	1287	CTCACCTCAGGCTTCCAGCCCGTTCAGCATCAAGCCCGACGAAGCAGACAACCCCAATCTG	1346
	D _b	1198	CTCACCTCGGCTTCCAGCGCTTCAGCATCAAGCCCGACGAAGCAGACAACCCCACTG	1257
	Q _y	1347	CTCACCTCAGCGGACCTCAGACACACC---ACCACTTACAGGCGAGTCTCTGTCCC	1403
	D _b	1258	CTCACCTCAGCGGACCTCAGCACCACTACCACTACCAAGGCGAGTCTATGTGTG	1317
	Q _y	1404	CGGAGGATGGCCCGACGCCCAAGTTCAGCTCACCAGTGGGACCTGCTCAGCCCCCTG	1463
	D _b	1318	AGGCAGGATGAGCCAGCGCCCAAGTTCAGCTCTCTAATGTCTACTGTCTCAGCCCACTG	1377
	Q _y	1464	GGTGGCGCCGCCACACACTGCACACAGCTCTCCCACTCTGAGGCCGAGGAGTTCGCTC	1523
	D _b	1378	GGGAGTGGCGCCATACTGTGCAACACAGCTCACCCACTCTGAGGCTGAGGACTTCGTC	1437
	Q _y	1524	TCCCGCTCTCCACCGAGAACTATTCGCTTCCTTGGCCCGAGCGCACGCAACATGACC	1583
	D _b	1438	TCCCGCTCTCCACCGAAAACTACTTTCTGTTCCCTTGGCCCGCGCACCGCAACATGGCC	1497
	Q _y	1584	TATGGGACTTCAACTTCTCGGGGCGGCTGATGATCCCTAATACAGGTATCAGCGCTC	1643
	D _b	1498	TACGGGACTTCAACTTCTCGGGGCGGCTGATGATCCCTAATACGGGATCAGGCTC	1557
	Q _y	1644	CTCATCCCCCGAGATGCCATACCCCGAGGGAGATCTATAGATCTTACCTCACGCTGCAC	1703

1558	DB	CTCATACCCCGGANTGCCATCCCCGAGGAAAGATCTACGAGATCTACTCCTCACACTGCAC	1611
1704	QY	AAGCCGGAAGACGTGAGGTTGCCCTAGCTGGCTGTGCAGACCTGTCTGAGTCCCACTGTT	1763
1618	DB	AAGCCAGAACGTGAGTTGCCCTAGCTGGCTGTGCAGACCTGTCTGAGTCCCACTGTT	1677
1764	QY	AGCTGTGGACCCCTGGCGTCTGTCTACCCGGCCAGTCACTCTGTGCTATGGACCACTGT	1823
1678	DB	AGCTGTGGGCCCCAGAGGTCCTGCTCACCCGGCCAGTCACTCTGTGCTCAATGGACCACTGT	1737
1824	QY	GGGAGCCCAAGCCCTGACAGCTGGAGCCTCGCGCTCAAAAAGCAGTCGTGCGAGGCGACG	1883
1738	DB	GGAGAGCCCAAGCCCTGACAGCTGAGTCTGCGCTCAAAAAGCAGTCGTGCGAGGCGAGT	1797
1884	QY	TGGAGCAGGAGTGTGCTGTCACCTGGGCGAGAGCGCCCTCCACCTCTACTACTGCCAG	1943
1798	DB	TGGG--AGGATGTGCTGCACCTTGGTGAGAGTCACCTTCCACCTCTACTACTGCCAG	1854
1944	QY	CTGAGGCGAGTGCCTGCTAGCTTTACCGAGAGCTGGGCGCGCTTGGCCCTGGTGGGA	2003
1855	DB	CTGAGGCGCGGGGCTGCTATGTCTTACCGAGCAGCTGGGCGCGCTTGGCCCTGGTAGGA	1914
2004	QY	GAGCCCTCAGCGTGGCTGCGCCCAAGCGCTCAAGCTGCTTCTGTGTTGGCGCGGTGGCC	2063
1915	DB	GAGCCCTCAGCGTGGCTGCGCCCAAGCGCTCAGGCTCTCTGTGTTGCTCCGTGGCC	1974
2064	QY	TGCACCTCCCTCGAGTAAACATCCGGGTCTACTGCTGTCATGACACCCACGATGCATC	2123
1975	DB	TGTAGTCCCTTGAGTACACATCCGAGTGTACTGCTTACAGCACCCACGACGCTCTC	2034
2124	QY	AAGAGGTGTGTGACGTGTGAGAACAGCTGGGGGGAAGCTGATCCAGGAGCCACGGTCT	2183
2035	DB	AAGAGGTGTGTGACGTGTGAGAACAGCTGAGTGTGGAAGCTGATCCAGGAGCCTCGCTC	2094
2184	QY	CTGCATCTCAGGACAGTTACCACAACTGGCGCTATCCATCCAGGATGTGCCAGCTCC	2243
2095	DB	CTGCATCTCAAGACAGTTACCAAACTAGCTCTCCATCCAGGAGTGCACGCTCC	2154
2244	QY	CTGTGGAGAGTAAGCTCCTTGTGACGTACAGGAGATCCCGCTTTTATCATCTGGAAT	2303
2155	DB	CTGTGGAGAGCAGCTACTTGTGACCTACGAGAGATCCCTTTTACCACTCTGGAAC	2214
2304	QY	GGCAGCAGCGGTACTTGCACTGCACTTCACTTGGAGCGTGTACGCCCCAGCACTAGT	2363
2215	DB	GGCAGCCAGCAGTACTTGCACTGCACTTCACTTGGAGCGTGTACGCCCCAGCACTAGT	2274
2364	QY	GACCTGGCTCAGAGCTGTGGTGTGCGAGTGGAGGCGGACGGGAGAGCTTCAGCATC	2423
2275	DB	GACCTGGCTCAGAGGTGTGGTGTGCGAGTGGAGGAGATGGGAGAGCTTCACATC	2334
2424	QY	AACTTCAACATCACCAAGACACAAGTTTGTCTGAGCTGTGGCTCTGGAGAGTGAAGCG	2483
2335	DB	AACTTCAACATCACTAAGACACAAGTTTGTCTGAAATTTGTGCTCTGGAGAGTGAAGGG	2394
2484	QY	GGGFTCCAGCCCTGGTGGGCCCCAGTGCCTTCAAGATTCGCCCTTCTCATTTGGCAGAAG	2543
2395	DB	GGGFTCCAGCCCTGGTGGGCCCCAGTGCCTTCAAGATTCGCCCTTCTCATTTGGCAGAAG	2454
2544	QY	ATAATTTCCAGCTGGACCACTGTAGAGGGGTGCCAGTGGCGGACTCTGGGCCAG	2603
2455	DB	ATCATCGCCAGTCTGGACCCCACTTGCAGCGGGGCGGCGGAGCTTGAAGACTCTAGCCCAAG	2514
2604	QY	AAACTCCACCTGGACAGCCATCTCAGCTTCTTTGCCCTCCAAAGCCAGCCCAAGCCATG	2663
2515	DB	AAACTCCACCTGGACAGCCATCTTAGCTTCTTTGCCCTCCAAAGCCAGCCCTTACAGCCATG	2574
2664	QY	ATCTCTCAACCTGTGGAGGGCGGGAATTCCTCCAAAGCCAGCCCTTACAGCCAGCTGGCTGCA	2723
2575	DB	ATCTCTCAACCTGTGGAGGGCGGGAATTCCTCCAAAGCCAGCCCTTACAGCCAGCTGGCTGCA	2634
2724	QY	GCAGTGGCTGAGCTGGGCGGCGGACAGCTGCGCTCTTTCAGTGTGCGAGGCTGAGTGC	2783
2635	DB	GCTGTGGCGGAGCTGGGCGGCGGACAGCTGCGCTCTTTCAGTGTGCGAGGCTGAGTGT	2694

Qy	2430	AAATACCAAGGACACAAGGTTTCTGAGCTGCTCTGAGAGTGAAGCGGGTC	2489
Db	1312	AAATACCAAGGACACAAGGTTTCTGAGCTGCTCTGAGAGTGAAGCGGGTC	1371
Qy	2490	CCAGCCCTGGTGGGCCCCAGTGCCTTCAAGATCCCTTCTCATTTGGCAGAGAATAATT	2549
Db	1372	CCAGCCCTGGTGGGCCCCAGTGCCTTCAAGATCCCTTCTCATTTGGCAGAGAATAATT	1431
Qy	2550	TCCAGCCTGGACCCACACCTGTAGGCGGGGTGCGACTGGCGGACTCTGGGCCAGAAATC	2609
Db	1432	TCCAGCCTGGACCCACACCTGTAGGCGGGGTGCGACTGGCGGACTCTGGGCCAGAAATC	1491
Qy	2610	CACCTGGACAGCAATCTCAGCTTCTTTGCCCTCAAGCCGAGCCCCAGACCATGATCCTC	2669
Db	1492	CACCTGGACAGCAATCTCAGCTTCTTTGCCCTCAAGCCGAGCCCCAGACCATGATCCTC	1551
Qy	2670	AACCTGTGGAGGCGCGGCACATTCGCCAAAGCGCAACCTCAGCCAGCTGGCTGCAGCAGTG	2729
Db	1552	AACCTGTGGAGGCGCGGCACATTCGCCAAAGCGCAACCTCAGCCAGCTGGCTGCAGCAGTG	1611
Qy	2730	GCTGGACTGGGCCAGCAGACGCTGGCCCTC-TTCACAGTG-TCCGAGGCTGAGTGCTGAG	2787
Db	1612	GCTGGGACTGGGCCAGCAGACGCTGGCTTCTTTTACAGTGTTCCGAGGCTGAGTGCTGAG	1671
Qy	2788	GCCGGCCAGGCCCGACACCTACACTCTCACCAGCTTTTGGCA--CCCAACAGGACAGGCA	2845
Db	1672	GCCGGCCAGG--CGAACACTACAAATTTTACAGTTTTTGGGAACCCACCAGGACAGCA	1729
Qy	2846	GAAGCCGGACAGGGGCCCTTCCCCACACCGGGGAGA	2881
Db	1730	GAAGCCGGACAGGGGCTTTTCCCAAACCGGGGAGA	1765

RESULT 9

US-10-256-702-2
; Sequence 2, Application US/10256702
; Publication No. US20030059859A1
; GENERAL INFORMATION:
; APPLICANT: Tessier-Lavigne, Marc
; Leonardo, E. David
; Hink, Lindsay
; Masu, Masayuki
; Kazuko, Keino-Masu
; TITLE OF INVENTION: Netrin Receptors
; NUMBER OF SEQUENCES: 8
; CORRESPONDENCE ADDRESS:
; ADDRESSER: SCIENCE & TECHNOLOGY LAW GROUP
; STREET: 268 BUSH STREET, SUITE 3200
; CITY: SAN FRANCISCO
; STATE: CALIFORNIA
; COUNTRY: USA
; ZIP: 94104
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patentin Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/10/256,702
; FILING DATE: 27-Sep-2002
; CLASSIFICATION: <Unknown>
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US/09/933,261
; FILING DATE: 20-Aug-2001
; APPLICATION NUMBER: 08/808,982
; FILING DATE: <Unknown>
; ATTORNEY/AGENT INFORMATION:
; NAME: OSMAN, RICHARD A
; REGISTRATION NUMBER: 36,627
; REFERENCE/DOCKET NUMBER: UC96-217
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (415) 343-4341
; TELEFAX: (415) 343-4342

QY 2010 CTGAGCTGGCTCGCGCCAGCGCCTCAAGCTCTTCTGTTTGGCGGGTGGCTGCACC 2069
Db |||||
QY 895 CTGAGCTGGCTCGCGCCAGCGCCTCAAGCTCTTCTGTTTGGCGGGTGGCTGCACC 954
Db |||||
QY 2070 TCCTCGAGTACAACTCCGGGTCTACTGCTGCATGACACCCAGATGCACTCAAGGAG 2129
Db |||||
QY 955 TCCTCGAGTACAACTCCGGGTCTACTGCTGCATGACACCCAGATGCACTCAAGGAG 1014
QY 2130 GTGGTCAGCTGAGAGCAGCTGGGGGACAGCTGATCCAGAGGACCGGGTCTGGAC 2189
Db |||||
QY 1015 GTGGTCAGCTGAGAGCAGCTGGGGGACAGCTGATCCAGAGGACCGGGTCTGGAC 1074
QY 2190 TTCAAGGACAGTTTACCAACCTCGCCCTATCCATCCAGATGTGCCAGCTCCCTGTGG 2249
Db |||||
QY 1075 TT-AAGGACAGTTTACCAACCT--GCCCTATCATCCAGATGTGCCAGCTCCCTGTGG 1131
QY 2250 AAGAGTAAGCTCTTGTGCTACCAAGGAGATCCCTTTTATCACATCTGGAATGGCAAG 2309
Db |||||
QY 1132 AAGAGTAAGCTCTTGTGCTACCAAGGAGATCCCTTTTATCACATCTGGAATGGCAAG 1191
QY 2310 CAGCGTACTTGCACGTGACCTTACCCCTGGAGCGTGTGAGCGGACGCTAGTGAAGCTG 2369
Db |||||
QY 1192 CAGCGTACTTGCACGTGACCTTACCCCTGGAGCGTGTGAGCGGACGCTAGTGAAGCTG 1251
QY 2370 GCTGCAAGCTGTGGGTGTGGAGGTGGAGGCGACGGGACAGCTTTCAGCATCAACTTC 2429
Db |||||
QY 1252 GCTGCAAGCTGTGGGTGTGGAGGTGGAGGCGACGGGACAGCTTTCAGCATCAACTTC 1311
QY 2430 AACATCACAAGGACACAAGAGTTTGTGAGCTGTGGGTCTGGAGAGTGAAGCGGGGTC 2489
Db |||||
QY 1312 AACATCACAAGGACACAAGAGTTTGTGAGCTGTGGGTCTGGAGAGTGAAGCGGGGTC 1371
QY 2490 CCAGCCCTGTGGGCGGCGGCTTCCAGATCCCTTCCATTCATTCGGGAGAGATAATT 2549
Db |||||
QY 1372 CCAGCCCTGTGGGCGGCGGCTTCCAGATCCCTTCCATTCATTCGGGAGAGATAATT 1431
QY 2550 TCCAGCTGGACCCACCTGTAGGGGGGTGCGGACTGGCGGACTTGGCCCGAGAACTC 2609
Db |||||
QY 1432 TCCAGCTGGACCCACCTGTAGGGGGGTGCGGACTGGCGGACTTGGCCCGAGAACTC 1491
QY 2610 CACTGGACAGCATCTCAGCTTCTTTGCTTCAAGATCCCTTCCATTCATTCGGGAGAGATAATT 1551
Db |||||
QY 1492 CACTGGACAGCATCTCAGCTTCTTTGCTTCAAGATCCCTTCCATTCATTCGGGAGAGATAATT 1551
QY 2670 AACCTGTGGAGCGCGGCACTTCCCGACGCAAGCTTCCAGCCAGCCAGCCATGATCCTC 2729
Db |||||
QY 1552 AACCTGTGGAGCGCGGCACTTCCCGACGCAAGCTTCCAGCCAGCCAGCCATGATCCTC 1611
QY 2730 GCTGAGCTGGGCGAGCCAGACGCTGGCCCTC-TTCACAGTG-TCCGAGGCTGAGTCTGAG 2787
Db |||||
QY 1612 GCTGAGCTGGGCGAGCCAGACGCTGGCCCTC-TTCACAGTG-TCCGAGGCTGAGTCTGAG 1671
QY 2788 GCCGCGAGGCGGCACTTCCCGACGCAAGCTTCCAGCCAGCCAGCCATGATCCTC 2845
Db |||||
QY 1672 GCCGCGAGG--CGAACACTACAAATTTTACCAGTTTGGGAACCCACCAAGGACAGCA 1729
QY 2846 GAAGCGGACAGGCGGCTTCCCGACGCGGAGCA 2881
Db |||||
QY 1730 GAAGCGGACAGGCGCTTTCCTTCCCAAAACCGGGGAGA 1765

RESULT 10

US-10-296-115-365
; Sequence 365, Application US/10296115
; Publication No. US20040053248A1
; GENERAL INFORMATION:
; APPLICANT: Hyseq Inc
; TITLE OF INVENTION: No. US20040053248A1el Nucleic Acids and Polypeptides
; FILE REFERENCE: 784PCT
; CURRENT APPLICATION NUMBER: US/10/296,115
; CURRENT FILING DATE: 2002-11-18
; PRIOR APPLICATION NUMBER: US09/488,725

; PRIOR FILING DATE: 2000-01-21
; PRIOR APPLICATION NUMBER: US09/552,317
; PRIOR FILING DATE: 2000-04-25
; NUMBER OF SEQ ID NOS: 1478
; SEQ ID NO 365
; LENGTH: 1321
; TYPE: DNA
; ORGANISM: Homo sapiens
US-10-296-115-365

Query Match 41.4%; Score 1193.6; DB 13; Length 1321;
Best Local Similarity 97.8%; Pred. No. 2.6e-287;
Matches 1295; Conservative 0; Mismatches 19; Indels 10; Gaps 8;

QY 1476 CACACACTGCACACAGCTCTCCACCTCTGAGGCGGAGAGTTCCTCTCCCGCTCTCC 1535
Db 1 CACACACTGCACACAGCTCTCCACCTCTGAGGCGGAGAGTTCCTCTCCCGCTCTCC 60
QY 1536 ACCCAGAACTACTTCGGCTCCCTGCCCGGAGGACCCAGCAACATGACCTATGGGACCTTC 1595
Db 61 ACCCAGAACTACTTCGGCTCCCTGCCCGGAGGACCCAGCAACATGACCTATGGGACCTTC 120
QY 1596 AACTTCTCTCGGGGCGGCTGATGATCCCTAATACAGGTATCAGCCTCTCATCCCCCA 1655
Db 121 AACTTCTCTCGGGGCGGCTGATGATCCCTAATACAGGTATCAGCCTCTCATCCCCCA 180
QY 1656 GATGCCATACCCCGAGGGAAGATCTATGAGATCTACCTCACGCTGCACAAGCGGAGAGAC 1715
Db 181 GATGCCATACCCCGAGGGAAGATCTATGAGATCTACCTCACGCTGCACAAGCGGAGAGAC 240
QY 1716 GTGAGTTGCCCTAGCTGGCTGTGAGACCTGTGAGTCCCATCTGTTAGCTGTGAGACC 1775
Db 241 GTGAGTTGCCCTAGCTGGCTGTGAGACCTGTGAGTCCCATCTGTTAGCTGTGAGACC 300
QY 1776 CCT-GCGCTCTGCTCACCGGCGAGTCACTCCT-GGCTATGACCACTGT-GGGGAGGCC 1832
Db 301 CTTGGGCTCTGCTTACCGGCGAGTCACTCTGGGTATGAGCACTGTGGGGAGGCC 360
QY 1833 AGCCTTGACAGCT-GGAGCTTGCCTCAAAAGCAGTCTGCGAGGGCAGCTGGGAGCA 1891
Db 361 AGCCTTGACAGCTGGGAGCTTGCCTCAAAAGCAGTCTGCGAGGGCAGCTGGG--A 417
QY 1892 GGATGTGTGACCTGGGCGAGGAGCGCTCCCACTCTACTACTGCGAGCTGGAGC 1951
Db 418 GGATGTGTGACCTGGGCGAGGAGCGCTCCCACTCTACTACTGCGAGCTGGAGC 477
QY 1952 CAGTCCCTGCTAGCTTTCACCGAGCAGTGGGCGCTTTCCTGCTGGGAGAGGCCCT 2011
Db 478 CAGTCCCTGCTAGCTTTCACCGAGCAGTGGGCGCTTTCCTGCTGGGAGAGGCCCT 537
QY 2012 CAGCTGGCTCGCGCAAGCGCTCAAGCTCTTCTGTTGGCGGCTGGCTGCACCTC 2071
Db 538 CAGCTGGCTCGCGCAAGCGCTCAAGCTCTTCTGTTGGCGGCTGGCTGCACCTC 597
QY 2072 CCTCGAGTACAACTCCGGGTCTACTGCTGATGACACCCAGATGCACTCAAGAGGT 2131
Db 598 CCTCGAGTACAACTCTGCTGCTACTGCTGATGACACTCAGATGCACTCAAGAGT 657
QY 2132 GGTGAGCTGGAGAGCAGCTGGGGGACAGCTGATCCAGGAGCCAGGGTCTGCACTT 2191
Db 658 GGTGAGCTGGAGAGCAGCTGGGGGACAGCTGATCCAGGAGCCAGGGTCTGCACTT 717
QY 2192 CAAGGACAGTTACCAACCTGGCCCTATCCATCCAGATGTGCCAGCTCCCTGTGAA 2251
Db 718 CAAGGACAGTTACCAACCTGGCCCTATCCATCCAGATGTGCCAGCTCCCTGTGAA 777
QY 2252 GAGTAAGCTCTTGTGAGTACCAAGGAGATCCCTTTTATCACTGTGAAATGGACGCA 2311
Db 778 GAGTAAGCTCTTGTGAGTACCAAGGAGATCCCTTTTATCACTGTGAAATGGACGCA 837
QY 2312 GCGGTACTTGCACGTGACCTTCACTGGAGCGTGTGAGCCCGAGCACTAGTGAAGCTGGC 2371
Db 838 GCGGTACTTGCACGTGACCTTCACTGGAGCGTGTGAGCCCGAGCACTAGTGAAGCTGGC 897


```
QY 2372 CTGCAAGCTGTGGGTGTGGAGGTGGAGGGGACGCGGAGAGCTTACGATCAACTTCAA 2431
Db 898 CTGCAAGCTGTGGGTGTGGAGGTGGAGGGGACGCGGAGAGCTTACGATCAACTTCAA 957
QY 2432 CATCAACAGGACACAAGGTTTGTGAGCTGTGCTCTGGAGAGTGAAGCGGGGTCCC 2491
Db 958 CATCAACAGGACACAAGGTTTGTGAGCTGTGCTCTGGAGAGTGAAGCGGGGTCCC 1017
QY 2492 AGCCCTGTGGGCCCCAGTGCCTTCAAGATCCCTTCTCTCATTTGGCGAAGATAATTTC 2551
Db 1018 AGCCCTGTGGGCCCCAGTGCCTTCAAGATCCCTTCTCTCATTTGGCGAAGATAATTTC 1077
QY 2552 CAGCCTGAGACCAACCTGTAGCGGGGTGCGAGCTGCGGAGTCTGGGCCAGAACTCCA 2611
Db 1078 CAGCCTGAGACCAACCTGTAGCGGGGTGCGAGTCTGGGAGTCTGGGCCAGAACTCCA 1137
QY 2612 CCTGACAGCATTCTCAGCTTCTTTGCTTCCAAAGCCAGCCCCACAGCCATGATCTCAA 2671
Db 1138 CCTGACAGCATTCTCAGCTTCTTTGCTTCCAAAGCCAGCCCCACAGCCATGATCTCAA 1197
QY 2672 COTGTGGAGGCGCGCACTTCCCAACGGCAACCTCAGCCAGCTGGTGCAGAGTGGC 2731
Db 1198 COTGTGGAGGCGCGCACTTCCCAACGGCAACCTCAGCCAGCTGGTGCAGAGTGGC 1257
QY 2732 T-GGACTGGGCGAGCAGAGCTGGCTC-TTCACAGTG-TGGAGGCTGAGTGTGAGG 2788
Db 1258 TGGGACTGGGCGAGCAGAGCTGGCTTCTTTTACAGTGTTCGGAGGCTGAGTGTGAGG 1317
QY 2789 CCGG 2792
Db 1318 CCGG 1321

RESULT 11
US-09-833-381-1806
; Sequence 1806, Application US/09833381
; Patent No. US2002013209A1
; GENERAL INFORMATION:
; APPLICANT: Robison, Keith E.
; TITLE OF INVENTION: No. US2002013209A1el Nucleic Acid and Protein Homologs
; FILE REFERENCE: 5800-1139
; CURRENT APPLICATION NUMBER: US/09/833,381
; CURRENT FILING DATE: 2001-04-11
; PRIOR APPLICATION NUMBER: 09/516,448
; PRIOR FILING DATE: 2000-02-29
; NUMBER OF SEQ ID NOS: 2050
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 1806
; LENGTH: 1282
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (1)-(1282)
; OTHER INFORMATION: n = A,T,C or G
US-09-833-381-1806

Query Match 39.8%; Score 1147.8; DB 9; Length 1282;
Best Local Similarity 97.7%; Pred. No. 6.7e-276;
Matches 1259; Conservative 0; Mismatches 18; Indels 11; Gaps 9;
QY 1510 CCGAGAGTTCGTCTCCGCTCTCCACCCAGAACTTCCGCTCCCTCCCGGAGGCA 1569
Db 1 CCGAGAGTTCGTCTCCGCTCTCCACCCAGAACTTCCGCTCCCTCCCGGAGGCA 60
QY 1570 CCAGCAACATGACTATGGACCTTCAACTTCTCGGGGCGGCTGATGATCCCTAATA 1629
Db 61 CCAGCAACATGACTATGGACCTTCAACTTCTCGGGGCGGCTGATGATCCCTAATA 120
QY 1630 CAGGTATCAGCTTCTCATCCCCCAGATGCCATACCCCGAGGAGATCTATGAGATCT 1689
Db 121 CAGGAATCAGCTTCTCATCCCCCAGATGCCATACCCCGAGGAGATCTATGAGATCT 180
```

```
QY 1690 ACCTCAGCTGCACAAGCCGGAAGACGTGAGGTGGCCCTAGCTGGCTGTGAGACCTGCG 1749
Db 181 ACCTCAGCTGCACAAGCCGGAAGACGTGAGGTGGCCCTAGCTGGCTGTGAGACCTGCG 240
QY 1750 TGAGTCCCATCGTTAGCTGTGGAGCCCTT-GGCGTCTGTGCTACCCGGCGAGTCACTCT- 1807
Db 241 TGAGTCCCATCGTTAGCTGTGGAGCCCTTGGCGTCTGTGCTACCCGGCGAGTCACTCTG 300
QY 1808 GGCTATGACCACTGTGGGAGCCCGACGCTGACAGCT-GGAGCCTGCGCCTCAAAAAGC 1866
Db 301 GGCTATGACCACTGTGGGAGCCCGACGCTGACAGCTGGGAGCTGTGCGCTCAAAAAGC 360
QY 1867 AGTCTGTGAGGCGAGCTGGGAGCAGGATGTGCTCAGCTTGGCGAGGAGCGCCCTTCCC 1926
Db 361 AGTCTGTGAGGCGAGCTGGGAGGATGTGCT--TGCACTGGCGAGGAGCGCCCTCCC 418
QY 1927 ACCTCTACTTTCGAGCTGGAGGCGAGTGCCTGTCTAGTCTTCCACCGAGAGCTGGGCG 1986
Db 419 ACCTCTACTTTCGAGCTGGAGGCGAGTGCCTGTCTAGTCTTCCACCGAGAGCTGAGCG 478
QY 1987 GCTTTGCCCTGTGGGAGGCGCCCTCAGCTGGCTGCCCGCAAGCGCTCAAGCTGCTTC 2046
Db 479 GCTATGCCCTGTGGGAGGCGCCCTCAGCTGGCTGCCCGCAAGCGCTCAAGCTGCTTC 538
QY 2047 TGTTTGGCGCGGTGGCTGCACCTCCCTCGAGTACAACAATCCGGGTCTACTGCTGCATG 2106
Db 539 TGTTTGGCGCGGTGGCTGCACCTCCCTCGAGTACAACAATCTGCTTACTGCTGCATG 598
QY 2107 ACACCCAGATGACTCAAGGAGGTGGTGCAGCTGGAGAGCAGCTGGGGGACAGCTGA 2166
Db 599 ACACCTCAGATGACTCAACGCTAGTGGTGCAGCTGGAGAGCAGCTGGGGGACAGCTGA 658
QY 2167 TCCAGGAGCCAGGGTCTGCACTTTCAGGAGCAGTTACCACAACTCGCCCTATCATCC 2226
Db 659 TCCAGGAGCCAGGGTCTGCACTTTCAGGAGCAGTTACCACAACTCGCCCTATCATCC 718
QY 2227 AGCATGTGCCAGCTCCCTGTGGAAGAGTAAGCTCTCTGTGAGTACCAAGAGATCCCC 2286
Db 719 AGCATGTGCCAGCTCCCTGTGGAAGAGTAAGCTCTCTGTGAGTACCAAGAGATCCCC 778
QY 2287 TTTATCACATCTGGAATGCAAGCGGTACTTTGACCTGACCTTACCTTACCTTGGAGCGTG 2346
Db 779 TTTATCACATCTGGAATGCAAGCGGTACTTTGACCTGACCTTACCTTGGAGCGTG 838
QY 2347 TCAGCCCCAGCACTAGTGCCTGCAAGCTGTGGGTGTGGAGTGGAGGGGCAAG 2406
Db 839 TCAGCCCCAGCACTAGTGCCTGCAAGCTGTGGGTGTGGAGTGGAGGGGCAAG 898
QY 2407 GGCAGAGCTTTCAGCATCAACTTCAACATCAACAAGGACACAAGTTTGTGAGTCTGTG 2466
Db 899 GGCAGAGCTTTCAGCATCAACTTCAACATCAACAAGGACACAAGTTTGTGAGTCTGTG 958
QY 2467 CTCTGGAGAGTGAAGCGGGGTCCCGAGCCCTGGTGGGCCCCAGTGCCTTCAAGATCCCT 2526
Db 959 CTCTGGAGAGTGAAGCGGGGTCCCGAGCCCTGGTGGGCCCCAGTGCCTTCAAGATCCCT 1018
QY 2527 TCTCTATTCGGCAGAGATAA-TTTCAGCCTGACCCCACTTGTAGGGGGGTGGCGAC 2585
Db 1019 TCTCTATTCGGCAGAGATAA-TTTCAGCCTGACCCCACTTGTAGGGGGGTGGCGAC 1078
QY 2586 TGGCGG-ACCTGTGGCCCCAGAAATCCACCTGGAGACCCATCTCAGCTTCTTTTGGCTCCAA 2644
Db 1079 TGGCGGAATCTTGGCCCCAGAAATCCACCTGGAGACCCATCTCAGCTTCTTTTGGCTCCAA 1138
QY 2645 GCCCAGCCCCAGACCATGATCTCAACTGTGGGAGCGCGCACTTCCCCAACGGCAA 2704
Db 1139 GCCCAGCCCCAGACCATGATCTCAACTGTGGGAGCGCGCACTTCCCCAACGGCAA 1198
QY 2705 CCTCAGCAGCTGGCTGCAGAGTGGCTGGAGTGGGCGAGCCAGAGAGCTGGCTTCTTTCAC 2764
Db 1199 CCTCAGCAGCTGG-TGCAGCAGTGGCTGGAGTGGGG--AGCAGAGCTGG-CTCTTTCAC 1254
```

QY 2765 AGTGTGGAGGCTGAGTGTGAGGCGG 2792
Db 1255 AGTGTGGAGGCTGAGTGTGAGGCGG 1282

RESULT 12

US-10-087-684-1
; Sequence 1, Application US/10087684
; Publication No. US20040029116A1
; GENERAL INFORMATION:
; APPLICANT: Edinger, Shlomit R.
; APPLICANT: MacDougall, John R.
; APPLICANT: Millet, Isabelle
; APPLICANT: Ellerman, Karen
; APPLICANT: Stone, David J.
; APPLICANT: Grosse, William M.
; APPLICANT: Lepley, Denise M.
; APPLICANT: Rieger, Daniel K.
; APPLICANT: Burgess, Catherine E.
; APPLICANT: Casman, Stacie, J.
; APPLICANT: Spytek, Kimberly A.
; APPLICANT: Boldog, Ferenc L.
; APPLICANT: Li, Li
; APPLICANT: Padigaru, Muralidhara
; APPLICANT: Mishra, Vishnu
; APPLICANT: Shenoy, Suresh G.
; APPLICANT: Rastelli, Luca
; APPLICANT: Tchernev, Velizar T.
; APPLICANT: Vernet, Corine A.M.
; APPLICANT: Zerhusen, Bryan D.
; APPLICANT: Malyankar, Uriel M.
; APPLICANT: Guo, Xiaojia
; APPLICANT: Miller, Charles E.
; APPLICANT: Gangolli, Esha A.
; TITLE OF INVENTION: PROTEINS AND NUCLEIC ACIDS ENCODING SAME
; FILE REFERENCE: 21402-214 CIP
; CURRENT APPLICATION NUMBER: US/10/087,684
; CURRENT FILING DATE: 2003-03-10
; PRIOR APPLICATION NUMBER: 60/253,834
; PRIOR FILING DATE: 2000-11-29
; PRIOR APPLICATION NUMBER: 60/250,926
; PRIOR FILING DATE: 2000-11-30
; PRIOR APPLICATION NUMBER: 60/264,180
; PRIOR FILING DATE: 2001-01-25
; PRIOR APPLICATION NUMBER: 60/274,194
; PRIOR FILING DATE: 2001-03-08
; PRIOR APPLICATION NUMBER: 60/313,656
; PRIOR FILING DATE: 2001-08-20
; PRIOR APPLICATION NUMBER: 60/327,456
; PRIOR FILING DATE: 2001-10-05
; NUMBER OF SEQ ID NOS: 220
; SOFTWARE: CuraSeqList version 0.1
; SEQ ID NO 1
; LENGTH: 2860
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (59)..(2857)
US-10-087-684-1

Query Match 30.9%; Score 889; DB 13; Length 2860;
Best Local Similarity 60.7%; Pred. No. 2,1e-211;
Matches 1637; Conservative 0; Mismatches 960; Indels 102; Gaps 7;

QY 184 ACCAGTCCCTGGTCCCAACCGGACCTGCTTCCCACTTCTGTTGGAGCCCGAGGATG 243
Db 168 ACTCTCCCTGTCAGCGCCAGCAGAGCCGCTGCCCTACTTCTGTCAGGAGCCACAGGAGC 227
QY 244 TGATCATGTCAGAACAGCCAGTGTCTGTGTGTCAGAGCCGTGCTCCGCGCAGCAGAGA 303
Db 228 CCTACATTTGTGAAGAACAGGCTGTGGAGCTCCGCTGCCGCGCTTCCCGCCACACAGA 287

QY 304 TCTTCTTCAAGTCRAACGGGAGTGGTGGCCAGGTGGACCACTGATCGAGCGAGCA 363
Db 288 TCTACTTCAAGTCAACGGCGAGTGGTGGCCAGAACGACCACTGATCGAGCGAGGCC 347
QY 364 CAGACGGGAGCAGTGGTGGAGCCGACCATGAGGTGCGCATTAATGTCTCAAGCAGCAGG 423
Db 348 TGGATGAGGCCACCGGTCTGCGGGTGGCGAGGTGCAGATCGAGGTGTCCGGCAGCAGG 407
QY 424 TCGAAGAGTGTTCGGCTGGAGGAATACTGTGCCAGTCCGTGGGATGAGTCTCTCGG 483
Db 408 TGGAGGAGCTCTTGGCTGGAGGATTAATGTGGTCCAGTGGCGTGGCTGGAGTCCGCG 467
QY 484 GCACCAACAGAGTCAGAACGCTTACATCGCATAGCAGATTGCCAAGAACTTCGAGC 543
Db 468 GCACCAACAGAGTCGCCGAGCTACGTCCGATCGCTACCTGCGCAAGAACTTCGATC 527
QY 544 AGGAGCGCTGGCCAAAGAGGTGCTCCCTGGAGCAGGGCATCGTGTGCTGCTGCTCCAC 603
Db 528 AGGAGCTCTGGGCAAGAGGTGCCCTCGACCATGAGGTTCCTGTCAGTGGCGCGC 587
QY 604 CGGAGGCTATCCCTCCAGCGAGGTGAGTGGCTCCGGAACGAGGACCTGTGGACCCGT 663
Db 588 CGGAGGGGTGCTGTGGCGGAGTGGATGGCTCAAGAAATGAGGATGTCATCGACCCCA 647
QY 664 CCTGAGCCCAATGTATATACATCAGCGGAGCAGACCTGTGTGGTGGTGCACAGGCCCGC 723
Db 648 CCAGGACACCAACTTCTGCTCAGCATCGACCAACCTCATCTCCCGAGGCGCGC 707
QY 724 TTGCTGACACGGCCCAACTACCTCGGTGGCCAAAGAAATCTGTGCAGCTGCCCGCAGG 783
Db 708 TGTGCGACACTGCCAACTATACCTGGTGGCCAAAGAAATCTGTGCACCAACGCCGAGCA 767
QY 784 CTTCCGCTGTGTATCTGTACGTGAACGGTGGTGGTGTGACGTGGACCGAGTGTGCG 843
Db 768 CCACTGCCACCTCATCTGTACGTGAATGGCGGTGGTCCAGTGGGCGAGAGTGTGTAC 827
QY 844 TCTGAGCGCCAGCTGTGGCGCGCTGGCAGAAACGAGCGCGAGCTGCACCAACCGCG 903
Db 828 CTTGCTCCAAACCGCTGTGGCGAGGTGGCAGAGCCACCGGACCTGTCACCAACCGCG 887
QY 904 CGCTCTCAAACGGGGCGCTTCTGTGAGGGGCAAGATGTCTCATGACGCCACCGTCTCT 963
Db 888 CTCCACTCAACGGAGGGCGCTTCTCGAGGGCCAGGCAATCCAGAA---GACCGCTGCA 944
QY 964 CTCTGCTGTGTCTGTGAACCGCAGCTGGAGCCCGTGGAGCAAGTGGTGGCGCTGTGG 1023
Db 945 CCACCATCTGCCCGAGTGCATGGGGGTGGACGGAGTGGAGCAAGTGGTGCAGCTGCAGCA 1004
QY 1024 TGGACTGCACCCACTGGCGGAGCGGTGAGTGTCTGACCCAGCACCCCGCAACGAGGGG 1083
Db 1005 CTGAGTGTGCCCACTGGCGGTAGCCGCGAGTGCATGGCGCGCTCGTGGTGGCCATCTTC 1064
QY 1084 AGGAGTGCACGGGCACTGACCTGGACACCCCACTGTACCACTGACCTCTGTGTACACA 1143
Db 1065 GTGACTGCAGCGGAGCGCTGTCTGACTTAAGAACTGCACAGATGGGTGTGTCATGCAAC 1124
QY 1144 GTGCTTGGCCCTGAGGACGTGGCCCTCTATGTGGGCTCATGCGCGTGGCGCTGTGCG 1203
Db 1125 TGGAGGCTCAGGGGATGCGGCGCTGTATGCGGGCTCGTGGTGGCCATCTTCGTGGTGC 1184
QY 1204 TGGTCTGTGCTGCTGTCTCTCATCTCTGTTTATTGCGGGAAGAGAGGGGTGTGACT 1263
Db 1185 TGGCAATCTCATGGCGTGGGGTGGTGGTGTATCCCGCGCAACTGCGGTGACTTTCGACA 1244
QY 1264 CAGATGTGGTGTGCTGTCTTCCATT---CTCACTCAGGCTTCCAGCCGTTCAGATCAAGC 1320
Db 1245 CAGACATCACTGATCATCTGCTGCCCTGACTGGTGGTGTTCACCCCGTCAACTTTAGA 1304
QY 1321 CCAGCAAGCAGACAAACCCCACTGTGT-----CACCATCCAGCCGACCTCAGACCA 1374
Db 1305 CGGCAAGGCCCAGTAAACCCGCGAGTCTCTACACCCCTCTGTGCTCTCTGACCTGACAGCA 1364
QY 1375 CCACCCTTACCAGGGCAGTCTGTGCCCGGAGGATGGGCCCGACCCCAAGTTCACG 1434

Db 1365 GCGCGGCAATCTACCGCGGACCCGCTGATGCTCCAGAGCTCCACCGCAAAATCCCA 1424
Qy 1435 TCACCAATGGGCACTGCTCAGCCCC----- 1460
Db 1425 TGACCAACTCTCTCTGTGGACCCCTTACCCAGCCCTTAAGGTCAAGGTCTACAGCTCA 1484
Qy 1461 -----CTGGGTGGCGCGCC 1476
Db 1485 GCACACAGGGCTCTGGCGCAGGCCCTGGCAGATGGGGCTGACCTGTGGGGGTCTTGGCGC 1544
Qy 1477 ACACACTGCACACAGCTCTCCACTCTGAGGCGAGGAGTTCGTCTCCCGCTCTCCA 1536
Db 1545 CTGGCACAATACCTAGCAATTCGCGGGGACACCACTTCTGTGACCTGCGCAGGGCA 1604
Qy 1537 CCCGAACCTAC-----TTCCGCTCCCTGCCCGAGGCAACCAATGACCTATGGGA 1590
Db 1605 GCCTCGGTTCCAGCAGCTCTTGGGCTGCGCGAGACCCAGGAGCAGCGTCAGCGCA 1664
Qy 1591 CTTTCAACTTCTCGGGGCGCGCTGATGATCCCTAATACAGGTATCAGCCTCTCATCC 1650
Db 1665 CTTTGGCTGCTGGGTGGAGGCTCAGCATCCCGGCAACAGGGGTGAGCTTGTGGTG 1724
Qy 1651 CCCAGATGCCATACCCGAGGAAGATCTATGAGATCTACCTCAGCTGCAACAGCCGG 1710
Db 1725 CCAATGGAGCCATTCGCCAGGCAAGTCTACGAGATGATCTACTCATCAACAGGCAG 1784
Qy 1711 AAGAGTGAAGTGGCCCTAGCTGCTGTFCAGACCCCTGTGATGCCATCGTTAGCTGTG 1770
Db 1785 AAGATACCTGCGCTTTTCAGAAAGGACCCAGACAGTATGAGCCCTCGGTGACCTGTG 1844
Qy 1771 GACCCCTGGGCTCTGTCTACCGGCGAGTCACTCTGCTATGACCACTGTGGGAGC 1830
Db 1845 GACCACAGGCTCTGTCTGTGCGCGCCGTATCCTCACCATGCCCCACTGTGCCGAAG 1904
Qy 1831 CAGCCCTGACAGCTGGAGCCTCGGCTCAGAAAGCAGTCGTGCGAGGCGAGCTGGAGC 1890
Db 1905 TCAGTCCCGTACGTGATCTTTCAGTCAAGACCCAGGCGCCACAGGGCCACTGGAG- 1963
Qy 1891 AGGATGTGTGCACTGGCGGAGGCGCCCTCCACCTCTACTACTGCTGCTGAGG 1950
Db 1964 --GAGTGTGACCTCTGATGAGGAGCCTGAACACACCTCTACTGCGAGCTGGAGC 2021
Qy 1951 CAGTGCCTGCTACGTCTTACCGAGCAGTGGCGCGCTTTCCTCTGGTGGGAGGCGCC 2010
Db 2022 CCAGGCGCTGTACATCTCTGTGAGCAGCTGGGACCTACGTTTACGGGGAGTCT 2081
Qy 2011 TCAGCTGTGCTGCGGCAAGCGCTCAAGCTGCTTCTGTTGGCGGCTGGCTGCACCT 2070
Db 2082 ATTCCGCTCAGAGTCAAGCGGCTCAGCTGGCGCTTTCGCGCCCGCCCTCTGACCT 2141
Qy 2071 CCTCGAGTACAACTCCGGGTCTACTGCTGATGACACCCACAGATGACTCAAGGAG 2130
Db 2142 CCTGGAGTACAGCTCCGGGTCTACTGCTGGAGGACAGCGCTGTAGCACTGAAGGAGG 2201
Qy 2131 TGGTGCAGTGGAGAGCAGCTGGGGGACAGCTGATCAGGAGCCACGGGTCTGCACT 2190
Db 2202 TGTGTGAGTGGAGCGGATCTTGGCGGATCTTGGTGGAGGAGCGCAACCGCTAATGT 2261
Qy 2191 TCAAGGACAGTACCAACCTCGGCTATCCATCCACGATGTCGCCAGCTCCCTGTGA 2250
Db 2262 TCNAGGACAGTACCAACCTCGGCTCTCCCTCCATGACCTCCCGCATTTGA 2321
Qy 2251 AGAGTAACTCTTGTAGCTACAGGAGATCCCTTTTATCATCTGGAATGGCAGC 2310
Db 2322 GGAGCAAGTGTCTGGGCAATATACAGGAGATCCCTTCTATCATCTTGGAGTGGAGCC 2381
Qy 2311 AGCGTACTTGCAGTCACTTCACTTCACTGAGGCTGTACGCCCGCAGCACTAGTGA 2370
Db 2382 AGAAGGCCCTCCACTGCACTTTCACCTTGGAGGAGCAGCTTGGCTCCACAGAGCTCA 2441
Qy 2371 CCTGCAAGCTGTGGGTGTGSCAGGTGGAGGCGACGGGCGAGCTTCAGCATCACTTCA 2430

Db 2442 COTGCAAGATCTGCTGGCGCAAGTGAAGGGAGGCGCAGATATTCAGCTGCATCCA 2501
Qy 2431 ACATCACCAAG---GACAAGAAGTTTGTGAGCTGTGGCTCTGAGAGTGAAGCGGGG 2487
Db 2502 CTCTGGCAGAGACACCTGCTGGCTCCCTGGACACTCTCTGTCTCTGCCCTGGCAGACTG 2561
Qy 2488 TCCAGGCTGCTGGGCGCCAGTGCCTTCAAGATCCCTTCTCATTCGGCAGAGAATAA 2547
Db 2562 TCACACCAAGCTGGACCTTATGCTTCAAGATCCCACTGTCCATCCGCCAGAAGATAT 2621
Qy 2548 TTTCCAGCTGGACCCACCCCTGTAGGGGGGTCGACGTGGGGGACTGTGGCCCAAGAAC 2607
Db 2622 GCAACAGCTAGATCCCCCAACTCACGGGGCAATGACTGGCGGATGTTAGCACAAGC 2681
Qy 2608 TCCACTCGACAGCCATCTCAGCTTCTTTCCTCAAGCCAGCCAGCCAGCCATGATCC 2667
Db 2682 TCTCTATGACCGGTACCTGAACTTTGCCACCAAGAGGAGCCCGGTGTGATCC 2741
Qy 2668 TCAACCTGTGGAGGCGGCGCACTTCCCCAAGCGCAACCTCAGCCAGCTGGCTGCAGCAG 2727
Db 2742 TGGACTCTGGGAAGCTCTGCAGCAGGAGGAGTGGGAGCTCAACAGCTTGGCGAGTGCCT 2801
Qy 2728 TGGCTGGACTGGGCCAGCCAGACCGTGGCTTCTTCAAGTGTGGAGGCTGAGTGTGA 2786
Db 2802 TGGAGGAGATGGGCAAGAGTGAATGCTGTGGCTGTGGCCACCGACCGGGGACTGCTGA 2860

RESULT 13

US-10-218-779-1

; Sequence 1, Application US/10218779
; Publication No. US2004002922A1

GENERAL INFORMATION:

; APPLICANT: Edinger, Shlomit
; APPLICANT: Machougall, John
; APPLICANT: Millet, Isabelle
; APPLICANT: Ellerman, Karen
; APPLICANT: Stone, David
; APPLICANT: Gerlach, Valerie
; APPLICANT: Grosse, William
; APPLICANT: Alsbrook II, John
; APPLICANT: Lepley, Denise
; APPLICANT: Rieger, Daniel
; APPLICANT: Burgess, Catherine
; APPLICANT: Casman, Stacie
; APPLICANT: Spytek, Kimberly
; APPLICANT: Boldog, Ferenc
; APPLICANT: Li, Li
; APPLICANT: Padigaru, Muralidhara
; APPLICANT: Mishra, Vishnu
; APPLICANT: Patturajan, Meera
; APPLICANT: Shenoy, Suresh
; APPLICANT: Rastelli, Luca
; APPLICANT: Tchernev, Velizar
; APPLICANT: Vernet, Corine
; APPLICANT: Zerhusen, Bryan
; APPLICANT: Malyankar, Uriel
; APPLICANT: Guo, Xiaojia
; APPLICANT: Miller, Charles
; APPLICANT: Gangoli, Bsha
; TITLE OF INVENTION: Proteins and Nucleic Acids Encoding Same
; FILE REFERENCE: 21402-214
; CURRENT APPLICATION NUMBER: US/10/218,779
; CURRENT FILING DATE: 2002-08-14
; PRIOR APPLICATION NUMBER: 60/253,834
; PRIOR FILING DATE: 2000-11-29
; PRIOR APPLICATION NUMBER: 60/250,-926
; PRIOR FILING DATE: 2000-11-30
; PRIOR APPLICATION NUMBER: 60/264,180
; PRIOR FILING DATE: 2001-01-25
; PRIOR APPLICATION NUMBER: 60/313,656
; PRIOR FILING DATE: 2001-08-20
; PRIOR APPLICATION NUMBER: 60/327,456
; PRIOR FILING DATE: 2001-10-05

```

; NUMBER OF SEQ ID NOS: 216
; SOFTWARE: Patentin Ver. 2.1
; SEQ ID NO 1
; LENGTH: 2860
; TYPE: DNA
; ORGANISM: Homo sapiens
US-10-218-779-1

```

Query Match 30.9%; Score 889; DB 13; Length 2860;
Best Local Similarity 60.7%; Pred. No. 2.1e-211;
Matches 1637; Conservative 0; Mismatches 960; Indels 102; Gaps 7;

QY	184	ACCCAGTGGCTGGTGCCAAACCGGACCTGCTTCCCACTTCTCTGTGAGGCCGAGGATG	243
DB	168	ACTCTTCCCGTCAGCGCCAGCAGAGCCGTGTCCTACTTCTCGAGGACCAAGGACG	227
QY	244	TGTACATCGTCAAGAACAGCAGTCGTGCTTGTTGTGCAAGGCCGTGTCGCGCACGACGA	303
DB	228	CCTACATTGTGAAGAACAAAGCCTGTGGAAGCTCCGTGCGCGCTTCCCGGCCACACAGA	287
QY	304	TCTTCTTCAAGTGCAAACGGGAGTGGGTGCGCCAGGTGGACCAAGTGTATCGAGCGCACA	363
DB	288	TCTACTTCAAGTGCAAACGGCGAGTGGGTACGCCAAGACGACCAAGTGTACACAGGAAGCC	347
QY	364	CAGACGGAGCAGTGGTGAAGCGACCAATGAGAGTCCGCATTAAATGCTCAAGAGCAGCAGG	423
DB	348	TGGATAGGCCACCGGTCTGCGGGTGCAGAGTGCAGATCGAGTGTGCGGCGCAGCAGG	407
QY	424	TCGAGAGGTCTTCGGGCTGGAGNATACTGGTGCAGTGGTGGCATGGAAGTCTCTCGG	483
DB	408	TGGAGGAGCTCTTTGGGCTGAGGAGTAACTGGTGCAGTGGGTGGCTTGGAGCTTCGCGG	467
QY	484	GCACCAACAGAGTCAGAAGGCCATATCCGCATAGCCAGATTGGCGAAGAACTTCGAGC	543
DB	468	GCACCAACAGAGTCGCCGAGCCTACGTCCGATCGCCTACTTGGCGAAGAACTTCGATC	527
QY	544	AGGAGCGCTGGCCAAAGAGGTGTCCTGGAGCAGGGCATCGTGTGCCCTGCGCGTCCAC	603
DB	528	AGGAGCCTCTGGCAAGAGAGTGCCCTGGACCATGAGGTTCTCTGCAAGTGCCCGCCG	587
QY	604	CGGAGGCGATCCCTCCAGCCGAGGTGAGTGGCTCCGGAACGAGGACCTGTTGGAACCGGT	663
DB	588	CGGAGGGGTCCTGTGGCCGAGGTGGAATGGCTCAAGAAATGAGGATGTTCATCGACCCCA	647
QY	664	CCCTGGACCCCAATGTATACATACGCGGGAGCACAGCCTGGTGGTGCACAGGCCCGCC	723
DB	648	CCCAGGACACCACTTCTCTGCTACCATCGACACACCTCATATCGCCGAGCGCCGCC	707
QY	724	TTGCTGACACGGCCCAACTACCTGGTGGCCAAAGAAACATCGTGGCACGTGCGCGCAGCG	783
DB	708	TGTCGACACTGCCCACTATACCTGCGTGGCCAAAGAAACATCGTGGCCAAACCGCGGACGA	767
QY	784	CCTCCGCTGTGTCATCGTCTACGTGAACGTGGGTGTCGACGTGACCGAGTGTGTCGG	843
DB	768	CCACTGCCACCGTCATCGTCTACGTGAATGGCGGCTGGTCCAGCTGGCGAGTGTGTCTAC	827
QY	844	TCTGACGCGCAGCTGTGGGCGCGCTGGCAGAAACGAGCCGAGCTGCACCAACCCCG	903
DB	828	CCTGCTCAACCGCTGTGGCGAGCTGGCAGAAAGCGCACCGGACCTGCACCAACCCCG	887
QY	904	CGGCTTCAACGGGGCGCTTTTGTGAGGGGCAGAAATGTCCATGACCGCACCGTCTCTCT	963
DB	888	CTCCACTCAACGGAGGGGCTTCTGTGAGGGCCAGGCATTCAGAA--GACCGCTTCCA	944
QY	964	CTCTGCTTGTCTCTGTGGAACGACGTGAGACCGGTGGAGCAAGTGTGTGGCTGTGGGG	1023
DB	948	CCACCATCTGCCCAAGTGCATGTGGGCGGTGGACGGAGTGGACAAGTGTGTGAGCTTCAGCA	1004
QY	1024	TGGAAGTGCACCCACTGGCGGAGCCGTGAGTGTCTGTGACCCAGCACCCCGCAACGGAGGGG	1083
DB	1008	CTGAGTGTGCCACTGCGGTAGCCGAGTGTGATGGCGCCGCCACCCACCAAGACGGAGGCC	1064
QY	1084	AGGAGTCCAGGGCACTGACCTGGACACCCGCAACTGTACAGTGAACCTCTGTGTACACA	1143

Db	2142	CCCTGGAGTACAGCCTCCGGGTCTACTGCTGGAGGACACGCCTGTATGACTCTGAAGGAGG	2201
Qy	2131	TGCTGCAGCTGGAGAAGCAGCTGGGGGACAGCTGATCCAGGAGCCACGGGTCTCTGCAC	2190
Db	2202	TGCTGGAGCTGGAGCGGACTCTGGCGGATACTTGGTGGAGGAGCCGAACCCGCTAATGT	2261
Qy	2191	TCAAGGACAGTTACACAAACCTGGCCCTATCCATCCACGATGTGCCCCAGCTCCCTGTGGA	2250
Db	2262	TCAAGGACAGTTAGTACAAACCTGGCGCTCTCCCTCCATGACTCCCCCATGCCCCATGGA	2321
Qy	2251	AGATGTAGCTCTTGTGAGCTACCAAGAGATCCCTTTTATCATCTGGAATGGCAGCG	2310
Db	2322	GGAGCAAGCTCTGGCCAAATACACAGAGATCCCCTTCTATCATTTGAGTGGCAGCC	2381
Qy	2311	AGCGGTACTTTCACCTGCACCTTACCCCTGGAGCGTGTACGCCCCAGCACTAGTGA	2370
Db	2382	AGAAGGCCCTCCACTGCACTTTACCCCTGGAGAGGCACAGCTTGGCCTCCACAGAGCTCA	2441
Qy	2371	CCTGCAGCTGTGGGTGTGGCAGCTGGAGGCGCAGGCGAGAGCTTCAGCATCAACTTCA	2430
Db	2442	CCTGCAAGATCTGGTGGCGGCAAGTGGAAAGGGAGGCGCAGATATTCAGCTGCATACCA	2501
Qy	2431	ACATCAACCAAG---GACACAAAGTTTGTGTGAGCTGCTGGCTCTCGAGAGTGAACGGGGG	2487
Db	2502	CTCTGGCAGAGACACCTGTGGCTCCCTGGACACTCTCTGCTCTGCCCTGGCAGCACTG	2561
Qy	2488	TCCAGACCTGGTGGGCCCCAGTGCCTCAAGATCCCTTCTCATTTGGCAGAGATAA	2547
Db	2562	TCACCAACCACTGGGACCTTATGCTTCAAGATCCCACTGTCCATCGCCAGAAATAT	2621
Qy	2548	TTTCCAGCCTGGACCCACCTGTAGCGGGGTGCCGACTGCGCGGACTCTGGCCCCAGAAAC	2607
Db	2622	GCAACAGCTAGATGCCCCCAACTCACGGGGCAATGACTGCGGATGTTAGCACAGAAGC	2681
Qy	2608	TCCACCTGGACAGCCATCTAGCTTCTTTGGCTTCAGCCCAGGCCCCACAGCCATGATCC	2667
Db	2682	TCTCTATGACCGGTACCTGAAATTACTTTGGCACCAAGCGAGCCCCACGGGTGTGATCC	2741
Qy	2668	TCAACCTGTGGAGCGGGCACTTCCCCAACGGCAACTCAGCCAGCTGGCTCGACGAG	2727
Db	2742	TGGACCTCTGGAGACTCTGCAGCAGCATGGGGACCTCAACAGCTTGGCGAGTGCCT	2801
Qy	2728	TGGCTGGACTGGGCCAGCCAGACGCTGGCCTTCTTCAAGTGTGGAGGCTGAGTGCTGA	2786
Db	2802	TGGAGGACATGGCAAGAGTGAATGTGGTGTGGCTGTGGCCACCAACCGACGGGACTGCTGA	2860

RESULT 14

US-10-087-684-3
; Sequence 3, Application US/10087684
; Publication NO. US20040029116A1
; GENERAL INFORMATION:
; APPLICANT: Edinger, Shlomit R.
; APPLICANT: MacDougall, John R.
; APPLICANT: Millet, Isabelle
; APPLICANT: Ellerman, Karen
; APPLICANT: Stone, David J.
; APPLICANT: Grosse, William M.
; APPLICANT: Lepley, Denise M.
; APPLICANT: Rieger, Daniel K.
; APPLICANT: Burgess, Catherine E.
; APPLICANT: Casman, Stacie, J.
; APPLICANT: Spytek, Kimberly A.
; APPLICANT: Boldog, Ferenc L.
; APPLICANT: Li, Li
; APPLICANT: Padigaru, Muralidhara
; APPLICANT: Mishra, Vishnu
; APPLICANT: Shenoy, Suresh G.
; APPLICANT: Rastelli, Luca
; APPLICANT: Tchernev, Velizar T.
; APPLICANT: Vernet, Corine A.M.
; APPLICANT: Zerhusen, Bryan D.
; APPLICANT: Malyanekar, Uriel M.

QY 784 CTTCCGCTGCTGCTATCTGCTGAAACGGTGGGTGGTGGAGTGGACCGAGTGGTCCG 843
Db |||||
QY 768 CCACTGCCACCGTCTGCTGTAAGTGGCGGCTGGTCCAGCTGGGAGAGTGGTCAAC 827
Db |||||
QY 844 TCTGAGCGCCAGCTGTTGGCGCGGTGGCAGAAACGGAGCCGGAGTGCACCAACCCGG 903
Db |||||
QY 828 CTTGCTCCAAACCGCTTGGCGGAGGTGGCAGAAGCGCACCCGGAGCTGCACCAACCCCG 887
Db |||||
QY 904 CGCTCTCAACGGGGCGCTTCTGTGAGGGGAGAGTGTCTCATGACCGACCGCTCTCT 963
Db |||||
QY 888 CTCCTACTCAACGAGGGGCTTCTGGAGGGCCAGGCATTTCCAGAA---GACCGCTGCA 944
Db |||||
QY 964 CTTCTGCTTGTCTGTGGACGGCAGCTGGAGCCGCTGGAGCAAGTGGTGGCGCTGGGGC 1023
Db |||||
QY 945 CCACCATCTGCCAGTCTGATGGGGCTGGACGAGTGGAGCAAGTGGTCAAGCTGCAGCA 1004
Db |||||
QY 1024 TGGACTGACCACTGCGGAGCGGTGAGTGTCTGACCCAGCACCCCGCAACGGAGGGG 1083
Db |||||
QY 1005 CTGAGTGTGCCACTGGCGTAGCGCGAGTGCATGGCGCCGCCACCCAGAACGGAGGCC 1064
Db |||||
QY 1084 AGGAGTGCAGGSCACTGACCTGGACACCGCAACTGTACCAAGTACCTCTGTGTACACA 1143
Db |||||
QY 1065 GTGACTGACGGGACGCTGTGACTCTAAGAACTGCAAGATGGCTGTGCATGCAAC 1124
Db |||||
QY 1144 GTGCTTCTGCCCCGTGAGGACGTGGCCCTCTATGTGGGCTCTATGGCGCTGGCGCTGTGCC 1203
Db |||||
QY 1125 TGGAGGCTCAGGGATGCGGCGCTGATATGGGGCTCGTGGTGGCCATCTTGTGGTGG 1184
Db |||||
QY 1204 TGGTCTGTGCTGTCTGCTCTATCTGCTGTTATGGCGAAGAGGAGGGGTGGACT 1263
Db |||||
QY 1185 TGGCAATCTCATGGCGGTGGGGTGGTGTATCGCGCGCAACTGCCGTGACTTCGACA 1244
Db |||||
QY 1264 CAGATGTGGCTGACTGCTCCATT---CTCACCTCAGGCTTCAGCCCGTCAAGCATCAAGC 1320
Db |||||
QY 1245 CAGACATCACTGACTCACTGCTGCCCTGACTGGTGGTGTTCACCCCGCTCAACTTTAAGA 1304
Db |||||
QY 1321 CCAGAAAGCAGACAAACCCCACTCTGCT-----CACCATCCAGCGGACCTCAGACCA 1374
Db |||||
QY 1305 CGGCAAGCCCAAGTAAACCGGAGCTCTTACACCCCTCTGTGCTCTGACTGACGCA 1364
Db |||||
QY 1375 CCACCACTACAGGGGAGTCTGTGTCGCCGAGAGTGGGCCAGCCCAAGTTCCAGC 1434
Db |||||
QY 1365 GCGCGGATCTACCGCGGACCCGTTGTATGCGCTGCAGGACTCCACCGACAAATCCCA 1424
Db |||||
QY 1435 TCACCAATGGGACCTGTCTCAGCCCC-----CTGGGTGGCGCGCC 1460
Db |||||
QY 1425 TGACCAACTCTCTCTGTGTAACCCCTTACCCAGCCTTAAAGTCAAGGTCTACAGTCCA 1484
Db |||||
QY 1461 -----CTGGGTGGCGCGCC 1476
Db |||||
QY 1485 GCACCAAGGCTCTGGGCGAGCTGGCAGATGGGGTGAACCTGTGGGGTCTTGGCGC 1544
Db |||||
QY 1477 ACACACTGACACAGCTCTCCACTCTGAGCCGAGAGTGTCTCTCCGCGCTCTCCA 1536
Db |||||
QY 1545 CTGGCACAATCCCTAGCGATTTGCGCGGACACCCACTTCTGCACTGCGCAGCGCA 1604
Db |||||
QY 1537 CCCAGAACTAC-----TTCCGCTCTCCCGCGGAGGACAGCAACATGACCTATGGA 1590
Db |||||
QY 1605 GCCTCGGTCCAGCAGCTCTTGGCGCTGCCCCGAGACCCAGGAGAGCGGTCAAGCGCA 1664
Db |||||
QY 1591 CTTTCAACTCTCTGGGGCGGCTGTATGATCCCTTAACAGGTATACAGCTCTCTCATCC 1650
Db |||||
QY 1665 CTTTGGCTGTGCTGGTGGAGGCTCAGCATCCCGGACAGGGGTGAGTGTCTGGTGC 1724
Db |||||
QY 1651 CCCAGATGCCATACCCCGAGGAGATCTATGAGATCTACTCACCGTGCACAGCCGG 1710
Db |||||
QY 1725 CCAATGGAGCAATTTCCCGAGGCAAGTTCTACAGATGTATCTACTCATCAACAGGCGAG 1784
Db |||||
QY 1711 AAGAGGTGAGTGTGCCCTAGCTGTGAGACCTGCTGAGTCCCATGCTTGTAGCTGTG 1770
Db |||||
QY 1785 AAGTACCTCGGCTTTTCAGAAAGGAGCCAGACAGTATTGAGCCCTCTCGGTGACCTGTG 1844
Db |||||

RESULT 15
US-10-218-779-3
; Sequence 3, Application US/10218779
; Publication No. US20040029222A1

QY 1771 GACCCCTGGGCTCTGCTTCAACCGGCGCAGTCTATCTGGGTATGGAACCACTGTGGGAGC 1830
Db |||||
QY 1845 GACCCACAGGCTCTCTGCTGTGCGGCCGCTCATCTCTCAACATGCCCACTGTGCCAAG 1904
Db |||||
QY 1831 CCAGCCCTGACAGCTGCGCTCAAAAAGCAGTCTGCGAGGCGAGTGGGAGC 1890
Db |||||
QY 1905 TCACTGCGGCTGACTGGATCTTTCACTCAAGACCCAGGCCCCACAGGGCCACTGGGAG- 1963
Db |||||
QY 1891 AGGATGTGTGCACTTGGGCGAGGAGCGCCCTCCACCTCTACTACTGCGAGCTGGAGG 1950
Db |||||
QY 1964 --GAGTGTGTGACCTTGGATGAGGAGACCTGAAACACCTGCTACTGCTGAGCTGGAGC 2021
Db |||||
QY 1951 CCAGTGCCTGCTACTCTTCAACGAGCAGTGGGCGGCTTTGGCCCTGGTGGGAGAGGCC 2010
Db |||||
QY 2022 CCAGGCGCTGTCAACATCTCTGCTGGACAGCTGGGACCTACGTGTTTCACGGCGAGTCT 2081
Db |||||
QY 2011 TCAGCTGCTGCTGCGGCAAGCGCTCAAGCTGCTTCTGTTTGGCGGCTGGCTGCACT 2070
Db |||||
QY 2082 ATTCCGCTCAGCAGTCAAGCGGCTCCAGCTGGCGCTTTCGCCCCCGGCTCTGCACT 2141
Db |||||
QY 2071 CCCTGAGTACAACTTCCGGTCTACTGCTGTGATGACACCCACAGATGCACTCAAGAGG 2130
Db |||||
QY 2142 CCCTGAGTACAGCTCTCGGCTTACTGCTGGAGGACAGCTGTAGCACTGAAGAGG 2201
Db |||||
QY 2131 TGGTGCAGCTGAGAGCAGCTGGGGGACAGTGTATCCAGAGCCACCGGCTCTGCACT 2190
Db |||||
QY 2202 TCTGGAGCTGAGCGGACTCTGGCGGATCTTGGTGGAGGAGCGCAACCGCTAATGT 2261
Db |||||
QY 2191 TCAAGGACAGTTTACCAACCTCGGCTTATCCATCCAGATGTGCCAGCTCCCTGTGGA 2250
Db |||||
QY 2262 TCAAGGACAGTTTACCAACCTCGGCTTCTCTCATGAGCTCCCGATGCCCATTTGA 2321
Db |||||
QY 2251 AGAGTAACTCTTGTGCTACAGGAGATCCCTTTTATCACTCTGGAATGGCAGC 2310
Db |||||
QY 2322 GGAGAACTGTGCGCAAAATACAGGAGATCCCTTCTATCACTTTGGAGTGGCAGC 2381
Db |||||
QY 2311 AGCGTACTTGCAGCTGACCTTCACTTGGAGCTGTGAGCCCGAGCACTAGTACCTGG 2370
Db |||||
QY 2382 AGAAGCCCTCCACTGCACTTCACTTGGAGGACAGCTTGGCTCTCCAGAGCTCA 2441
Db |||||
QY 2371 CTTCAAGCTGTGGGTGGAGGTGGAGGCGAGGCGAGGCTTCAGACTCAACTCA 2430
Db |||||
QY 2442 CTTGCAAGATCTGCTGCGGCAAGTGGAGGAGGCGAGATATTCAGCTTGCATACA 2501
Db |||||
QY 2431 ACATCAACAG---GACAAAGTTTGTGCTGCTGCTCTGAGAGTGAAGCGGG 2487
Db |||||
QY 2502 CTCTGGCAGAGACACTGCTGGCTCCCTGGACACTCTCTGCTCTGCGCTTGCAGCACTG 2561
Db |||||
QY 2488 TCCAGAGCTGTGGGCGGCTTCAAGATCCCTTCTCTATTCGGGAGAAAGATAA 2547
Db |||||
QY 2562 TCACACCCAGCTGGGACCTTATGCTTCAAGATCCCACTGTCCATTCGCCAGAGATAT 2621
Db |||||
QY 2548 TTTCCAGCTGGACCCACTGTAGGGGGGTGGCGGCTTGGCGGACTCTGCGCCAGAAC 2607
Db |||||
QY 2622 GCAACAGCTAGATGCCCAACTCAACGGGCAATGACTGGCGGATGTTAGCACAGAAC 2681
Db |||||
QY 2608 TCCACTTGGACAGCACTCTAGCTTCTTTGCTCCAAAGCCAGCCCGCCACAGCCATGATCC 2667
Db |||||
QY 2682 TCTCTATGAGCGGCTTACCTGAATTTACTTTGCGCAAAAGCGAGCCCGCTGTATCC 2741
Db |||||
QY 2668 TCAACCTGTGGGAGGCGGCTTCCCAACGGCACTTCCAGCGAGCTTGGCTGGAGAG 2727
Db |||||
QY 2742 TGGACCTCTTGGAGAGCTTCTGCAAGAGCATGAGGAGCTTCAACAGCTTGGCGAGTGCCT 2801
Db |||||
QY 2728 TGGTGAAGTGGGCGAGCAGACCTGCGCTTCTTCAAGTGTCCGAGGCTGAGTGTGCTGA 2786
Db |||||
QY 2802 TGGAGGAGATGGGCAAGAGTGAATGCTGTGGCTGTGGCCACCGAGCGGAGTGTCTGA 2860
Db |||||

```

; GENERAL INFORMATION:
; APPLICANT: Edinger, Shlomit
; APPLICANT: MacDougall, John
; APPLICANT: Millet, Isabelle
; APPLICANT: Ellerman, Karen
; APPLICANT: Stone, David
; APPLICANT: Gerlach, Valerie
; APPLICANT: Grosse, William
; APPLICANT: Alsobrook II, John
; APPLICANT: Lepley, Denise
; APPLICANT: Rieger, Daniel
; APPLICANT: Burgess, Catherine
; APPLICANT: Casman, Stacie
; APPLICANT: Spytek, Kimberly
; APPLICANT: Boldog, Ferenc
; APPLICANT: Li, Li
; APPLICANT: Padigaru, Muralidhara
; APPLICANT: Mishra, Vishnu
; APPLICANT: Patturajan, Meera
; APPLICANT: Shenoy, Suresh
; APPLICANT: Rastelli, Luca
; APPLICANT: Tchernev, Velizar
; APPLICANT: Vernet, Corine
; APPLICANT: Zerhusen, Bryan
; APPLICANT: Malyankar, Uriel
; APPLICANT: Guo, Xiaojia
; APPLICANT: Miller, Charles
; APPLICANT: Gangolli, Bsha
; TITLE OF INVENTION: Proteins and Nucleic Acids Encoding Same
; FILE REFERENCE: 21402-214
; CURRENT APPLICATION NUMBER: US/10/218,779
; CURRENT FILING DATE: 2002-08-14
; PRIOR APPLICATION NUMBER: 60/253,834
; PRIOR FILING DATE: 2000-11-29
; PRIOR APPLICATION NUMBER: 60/250,-926
; PRIOR FILING DATE: 2000-11-30
; PRIOR APPLICATION NUMBER: 60/264,180
; PRIOR FILING DATE: 2001-01-25
; PRIOR APPLICATION NUMBER: 60/313,656
; PRIOR FILING DATE: 2001-08-20
; PRIOR APPLICATION NUMBER: 60/327,456
; PRIOR FILING DATE: 2001-10-05
; NUMBER OF SEQ ID NOS: 216
; SOFTWARE: Patentin Ver. 2.1
; SEQ ID NO 3
; LENGTH: 2860
; TYPE: DNA
; ORGANISM: Homo sapiens
; US-10-218-779-3

Query Match 30.8%; Score 887.4; DB 13; Length 2860;
Best Local Similarity 60.6%; Pred. No. 5.4e-211;
Matches 1636; Conservative 0; Mismatches 961; Indels 102; Gaps 7;

QY 184 ACCAGTGTGTCAGCAACCCGAGCTGCTTCCCACTTCCGTGGAGCCCGAGATG 243
Db 168 ACTCTTCCCGTCAGCGCCAGAGACCGCTGCGCTTCTCTCGAGAGCCACAGAG 227
QY 244 TGTACATCGTCAAGAAACAAGCCAGTGTGCTTGTGTGCAAGGCGGTGCCCGCCAGCA 303
Db 228 CCTACATTGTGAACACAGCTGTGGAGCTTCGCTTCCGCGCGCTTCCCGCCACACAGA 287
QY 304 TCTTCTTAAGTCAACGGGGAGTGGTGGCGCAGGTGGACCAAGTGTATCGAGCGAGCA 363
Db 288 TCTACTTCAAGTCAACGGGGAGTGGTGGCGCAGGACGACCAAGTGTATCGAGCGAGGCC 347
QY 364 CAGACGGGACAGTGTGAGCCAGCCAGCATGAGGTGCCCATTAATGTCTCAAGGCGAGG 423
Db 348 TGGATGAGCCACCGGCGCTGCGGTGCGAGGTGAGATCGAGGTGTCGCGGAGCAGG 407
QY 424 TCGAGAAGGTGTTTCGGGCTGGAGGAATACGTGGTSCCAAGTGTGGCATGAGTCTCTCG 483
Db 408 TGGAGGAGCTCTTTGGGCTGGAGGATTACTGGTSCCAAGTGTGGCTGGAGTCTCGAG 467

484 GCACCAACAAGTTCAGAGGCGCTATCCGATAGCCAGATTGCGAAGAACTTCGAGC 543
468 GCACCAACAAGTTCAGAGGCGCTATCCGATAGCCAGATTGCGAAGAACTTCGATC 527
544 AGGAGCGCTGCGCAAGGAGGTGTCCTCGAGCAGGCGATCGTGTGCGCTGCGCTCCG 603
528 AGGAGCGCTGCGCAAGGAGGTGTCCTCGAGCAGGCGATCGTGTGCGCTGCGCTCCG 587
604 CGGAGGCGATCCCTCCAGCGAGGTGAGTGGCTCGGAAACAGAGACCTGTGTGGACCG 663
588 CGGAGGCGGTGCTGTGCGCGAGGTGGAATGGCTCAAGAAATGAGGATGTTCATCGA 647
664 CCCTGGACCCCAATGTATACATCAGCGGAGCAGACGCTGTGTGGCAGAGCGCGCG 723
648 CCCAGGACACCAATTTCTGTGTCACTCAGCACCACCACTCATCATTCGCGAGGCG 707
724 TTGCTGACACGCGCAACTACACTGCTGTGGCGCAAGAAACATCGTGGCACTGCGCG 783
708 TGTGAGACACTGCGCAACTATACCTGGTGGCCCAAGAAACATCGTGGCAAAACGCG 767
784 CTTCCGCTGTGTCTATCGTGTACGTGAACGTTGGTGGTGTGCGACGTGGACCGAGT 843
768 CCACTGCGCACCGTCTATCGTGTACGTGAATGGCGGCTGTGTCCAGCTGGGCGAG 827
844 TCTGAGCGCGCAGCTGTGGCGCGCTGSCAGAAACGAGCGCGAGCTGCACCAACCCG 903
828 CTTGCTTCAACCGCTGTGGCGCGCTGSCAGAGCTGSCAGAGCGCACCGGACCTGCA 887
904 CGCTCTCTCAACGCGGCGCTTTCTGTGAGGGCGAGAAATGTCATGACCGACCGTCT 963
888 CTCCTACTCAACGAGGCGCTTTCTGCGAGGCGCAGGCATTCAGAA---GACGCGCT 944
964 CTCTGCTGTCTGTGAGCGCAGCTGGAGCGCTGGAGCAAGTGTGGCTGGCTGGGGC 1023
945 CCACCATCTGCCCGAGTCGATGGGCGTGGACGAGTGGAGCAAGTGTGAGCTGCGAG 1004
1024 TGGAGTGCACCCACTGGCGGAGCGTGTGAGTGTCTGTACCCAGCACCGCCCAACG 1083
1005 CTGAGTGTGCGCCACTGGCGGTAGCGGAGTGTATGGCGCCCGCACCCAGAGAGG 1064
1084 AGGAGTGCACGAGGCACTGACCTGGACACCCGCAACTGTATACAGTACCTGTGTG 1143
1065 GTGACTGACGCGGAGCGCTGTGCTGACTCTAAGAACTGCACAGATGGGCTGTG 1124
1144 GTGCTTGTGCGCTGAGGAGTGGCGCTCTATGTGGCGCTCATCGCGTGGCGCTG 1203
1125 TGGAGGCGCTCAGGGGATGCGGCGCTGTATGCGGGGCTCGTGTGGCCATCTTCT 1184
1204 TGTCTCTGTGTGTCTGTCTCTCATCTCTGTTTATTCGCGGAGAGAGGCGCTGG 1263
1185 TGGCAATCTCTATGGCGGTGGGGGTGTGTGTACCGCGCAACTCCCGTGTACTCG 1244
1264 CAGATGTGGCTGATCGTCCATT---CTCACTCAGGTTTCAGCGCGCTGAGCATCA 1320
1245 CAGACATCACTGACTCATCTGTGCGCTGACTGTGTGTTTCCACCGCGTCACTTT 1304
1321 CCAGCAAGAGACACACCCCGCATCTGCT-----CACCATCCAGCGCGACCTC 1374
1305 CGGCAAGGCGCCAGTAACCGCGAGCTCTACACCCCTCTGTGTCTCTCTGACTT 1364
1375 CCACCACTTACCAAGGCGAGTCTCTGTCCCGGAGGATGGGCGCCAGCCCCAAGT 1434
1365 GCGCGCGCATCTACCGGAGCCCGTGTATGCCCTCGAGGACTCCACCGCAAAAT 1424
1435 TCACCAATGGGCACTGTCTCAGCCCC----- 1460
1425 TGACCAACTCTCTCTGTGGACCCCTTACCACCGCTTAAGGTCAAGGTCTACAGT 1484
1461 -----CTGGGTGGGCGCGCGC 1476
1485 GCACCAAGGCGCTCTGGGCGAGGCGCTGGCAGATGGGGCTGACCTGCTGGGGGT 1544
```


Qy	1477	ACACACTGGACACACAGCTCTCCCACTCTGAGGCGGAGGATTGCTCTCCGCGCTCTCCA	1533
Db	1545	CTGGCACATACCTTAGCGAATTTGCGCCGGGACACCCATTCTCTGCACCTGCGCAGCGCA	1604
Qy	1537	CCAGAACTAC-----TTCCGCTCCCTGCGCCGAGGACACAGCAACATGACCTTATGGGA	1590
Db	1605	GCCTCGTTCCAGCAGCTCTTTGGCCTTGCCCGGAGCCAGAGGAGCAGGTTGAGCGGCA	1664
Qy	1591	CCTTCAACTTCTCTGGGGCGCGGTGATGCCCTTAATACAGGTATCAGACCTCTCTCATCC	1650
Db	1665	CCTTTGGCTGCTGGTGGAGGCTCAGCATCCCGGCAAGGGTGAGTCTGCTGCTGGTGC	1724
Qy	1651	CCCAGATGCCATACCCCGAGGGAAGATCTAAGAGATCTACTCACGCTGCACAAGCCGG	1710
Db	1725	CCAAITGGAGCAATTTCCCGAGGCAAGTTCTACGAGATGTATCTACTCATCAACAGGCAG	1784
Qy	1711	AAGACGTGAGGTGGCCCTAGCTGCTCAGACCCCTGCTGAGTCCCATGTTAGCTGTG	1770
Db	1785	AAAGTACCTGCGGTTTTCAGAGGGACCCAGACAGTATTGAGCCCTCGGTGACCTGTG	1844
Qy	1771	GACCCCTGGGCTCTGCTACCCGGCGAGTCATCCTGGCTATGGACCATGTGGGAGC	1830
Db	1845	GACCCACAGGCTCTCTGCTGTCGCGCCCGTCATCTCACATGCCCATGTGCCGAAG	1904
Qy	1831	CCAGCCTGACAGCTGGAGCTCGGCTCAAAAGCAGTCTGCGAGGCGAGCTGGGAGC	1890
Db	1905	TCAGTGCCTGTGACTGGATCTTTTCAGCTCAAGACCCAGGCCACCAGGGCCACTGGGAG-	1963
Qy	1891	AGGATGTGTCACCTGGCGGAGGAGCGCCCTCCACCTCTACTGCTACTGCGACTGGAGG	1950
Db	1964	--GAGGTGGTACCCCTGGATGAGGAGACCTGAACACACCCCTCTACTGCGACTGGAGC	2021
Qy	1951	CCAGTGCCTGTACGTTTCCAGCAGCAGCTGGCGCGCTTTGCCCCTGGTGGGAGAGGCC	2010
Db	2022	CCAGGCTGTGCATCTCTGCTGACAGCTGGCACCTAGCTGTTTCACGGGCGATCTCT	2081
Qy	2011	TCAGCTGGCTGGCGCCAGGGCTCAAGCTGTTCTGTTTGGCGCGGTGGCTGCACT	2070
Db	2082	ATTTCCCGCTCAGCAGTCAAGCGGCTCAGCTGGCGGCTTTGGCCCCCGCCCTCTGCACT	2141
Qy	2071	CCCTCGAGTACACATCCGGGTCTACTGCTGCATGACACCCAGATGCACTCAAGGAGG	2130
Db	2142	CCCTGGATACAGCTCCGGGTCTACTGCTGGAGCACGCTGTAGCACTGAAGGAGG	2201
Qy	2131	TGCTGACGTGGAGAAAGCAGCTGGGGGAGCAGCTGATCCAGAGGCCACGGTCTTGCACT	2190
Db	2202	TGCTGGAGCTGGAGCGGACTCTGGCGGATCTTGGTGGAGAGCGCAACCGCTAATGT	2261
Qy	2191	TCAGAGCATTTACCAACCTCGGCTATCCATCCACGATGTGCCAGCTCCCTGTGGA	2250
Db	2262	TCAAGACAGTTACCAACAACCTCGGCTCTCCCTCCATGACCTCCCCCATGCCATTGGA	2321
Qy	2251	AGAGTAAGCTCTTGTTCAGCTACAGGAGATCCCTTTTATCACATCTGGAATGGACGC	2310
Db	2322	GGAGCAAGCTGTGGCCAAATACAGGAGATCCCTTCTATCACATTTGGAGTGGCAGCC	2381
Qy	2311	AGCGGTATTGCACTGCACCTTCAACCCTGGAGCGGTCTCAGCCCCAGCACTAGTGAACCTGG	2370
Db	2382	AGAAGGCCCTCCACTGCACCTTTCACCCTGGAGGCGCAGCTTGGCTTCCACAGAGCTCA	2441
Qy	2371	CCTGCAAGCTGTGGGTGTGGCAGGTGGAGGGCGAGGGCAGAGCTTCAGCAATCACTTCA	2430
Db	2442	CTTGCAAGATCTGCGTGCAGCAAGTGGAGGGGAGGGCCAGATATTCAGAGCTGCATACCA	2501
Qy	2431	ACATCAACCAAG---GACACAAAGTTTGTCTGAGCTGCTGCTCTGAGAGTGAAGCGGGG	2487
Db	2502	CTCTGGCAGAGACACTGCTGGCTCCCTGGACACTCTCTGCTCTGCCCTGGCAGCACTG	2561
Qy	2488	TCCAGGCCCTGTGGGCCCGCAGTGCCTTCAAGATCCCTTTCCTCATTCGGCAGAGATAA	2547
Db	2562	TCACCAACCCAGCTGGACCTTATGCTTCAAGATCCACTGTCCATTCGCCCAAGAGATAT	2621
Qy	2548	TTTCCAGCTGGACCCACCTGTAGCGGGGTGCCGACTGGCGGACTCTGGCCAGCAAAAC	2607

This Page Blank (uspio)

GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: October 5, 2004, 15:09:57 ; Search time 7172 Seconds

(without alignments)

11995.670 Million cell updates/sec

Title: US-09-970-944-1

Perfect score: 2881

Sequence: 1 agctgggctcgggtgag.....ccttccacacccggggaga 2881

Scoring table: IDENTITY NUC

Gapop 10.0 , Gapext 1.0

Searched: 27513289 seqs, 14931090276 residues

Total number of hits satisfying chosen parameters: 55026578

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

EST:*

1:	em_estba:*
2:	em_esthum:*
3:	em_estin:*
4:	em_estmu:*
5:	em_estov:*
6:	em_estpl:*
7:	em_estro:*
8:	em_htc:*
9:	gb_est1:*
10:	gb_est2:*
11:	gb_htc:*
12:	gb_est3:*
13:	gb_est4:*
14:	gb_est5:*
15:	em_estfun:*
16:	em_estom:*
17:	em_gss_hum:*
18:	em_gss_inv:*
19:	em_gss_pln:*
20:	em_gss_vrt:*
21:	em_gss_fun:*
22:	em_gss_mam:*
23:	em_gss_mus:*
24:	em_gss_pro:*
25:	em_gss_rod:*
26:	em_gss_phg:*
27:	em_gss_vrl:*
28:	gb_gss1:*
29:	gb_gss2:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	907.8	31.5	2802	29 AY406491	AY406491 Homo sapi
2	898.6	31.2	2791	29 AY406493	AY406493 Mus muscu
3	871.6	30.3	3790	11 AK031655	AK031655 Mus muscu
4	818.4	28.4	3866	11 AK018177	AK018177 Mus muscu

5	763.2	26.5	2532	29 AY411747	AY411747 Homo sapi
6	763.2	26.5	2802	29 AY406492	AY406492 Pan trogl
7	756.2	26.2	1034	12 BI758231	BI758231 603029876
8	726.4	25.2	2532	29 AY411749	AY411749 Mus muscu
9	726.2	25.1	909	13 BX364574	BX364574 BX364574
10	722.2	25.1	788	9 AI951556	AI951556 wv36f04.x
11	711.2	24.7	1532	11 BC033727	BC033727 Homo sapi
12	710	24.6	818	12 BI818609	BI818609 603033362
13	709.8	24.6	843	13 BX348193	BX348193 BX348193
14	700.2	24.3	939	13 BX345406	BX345406 BX345406
15	679.4	23.6	1040	13 BX345407	BX345407 BX345407
16	664.2	23.1	1201	13 BX422753	BX422753 BX422753
17	664	23.0	889	13 BC0691915	BC0691915 AGENCOURT
18	653.6	22.7	934	10 BF311804	BF311804 601897316
19	648.6	22.5	2532	29 AY411748	AY411748 Pan trogl
20	646.4	22.4	859	10 BF311896	BF311896 601897733
21	644.2	22.4	874	13 BC0689148	BC0689148 AGENCOURT
22	604.4	21.0	756	13 BU612387	BU612387 UI-M-EWO-
23	596.4	20.7	662	9 AL516580	AL516580 AL516580
24	579	20.1	2775	29 AY401471	AY401471 Mus muscu
25	573.4	19.9	604	10 BE314370	BE314370 601147261
26	573.4	19.9	2775	29 AY401469	AY401469 Homo sapi
27	554.6	19.3	1175	10 BF530640	BF530640 602071931
28	548	19.0	788	14 CA317532	CA317532 UI-M-FW0-
29	539.4	18.7	678	14 CA749784	CA749784 UI-M-FD0-
30	524.4	18.2	824	12 BI737024	BI737024 603360874
31	493	17.1	1030	12 BM479478	BM479478 AGENCOURT
32	471.2	16.4	540	12 BI467799	BI467799 389583 MA
33	467.8	16.2	917	13 BX453456	BX453456 BX453456
34	463.6	16.1	1028	13 BX331575	BX331575 BX331575
35	449.8	15.6	987	13 BX452510	BX452510 BX452510
36	445	15.4	2507	29 AY408470	AY408470 Pan trogl
37	433	15.0	471	13 BX282095	BX282095 BX282095
38	430.6	14.9	675	14 CA315487	CA315487 UI-M-FW0-
39	424.4	14.7	843	12 BG913440	BG913440 602811321
40	424	14.7	780	12 BG470852	BG470852 602511705
41	416	14.4	718	14 CA318180	CA318180 UI-M-FW0-
42	415.4	14.4	1147	13 BU840446	BU840446 AGENCOURT
43	413	14.3	960	12 BG329812	BG329812 602429458
44	411.6	14.3	1111	12 BG298307	BG298307 602397080
45	411.4	14.3	749	14 CF735417	CF735417 UI-M-HB0-

ALIGNMENTS

RESULT 1	AY406491	2802 bp	DNA	linear	GSS 15-DEC-2003
LOCUS	AY406491	Homo sapiens	UNC5C gene, VIRTUAL TRANSCRIPT, partial sequence,		
DEFINITION	AY406491	Genomic survey sequence.			
ACCESSION	AY406491.1	GI:39762465			
VERSION	GSS.				
KEYWORDS	Homo sapiens (human)				
SOURCE	Homo sapiens				
ORGANISM	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.				
REFERENCE	1 (bases 1 to 2802)				
AUTHORS	Clark,A.G., Glanowski,S., Nielson,R., Thomas,P., Kejariwal,A., Todd,M.A., Tanenbaum,D.M., Civello,D.R., Lu,F., Murphy,B., Ferriera,S., Wang,G., Zheng,X.H., White,T.J., Shinsky,J.J., Adams,M.D. and Cargill,M.				
TITLE	Inferred nonneutral evolution from human-chimp-mouse orthologous gene trios				
JOURNAL	Science 302 (5652), 1960-1963 (2003)				
PUBMED	14671302				
REFERENCE	2 (bases 1 to 2802)				
AUTHORS	Clark,A.G., Glanowski,S., Nielson,R., Thomas,P., Kejariwal,A., Todd,M.A., Tanenbaum,D.M., Civello,D.R., Lu,F., Murphy,B., Ferriera,S., Wang,G., Zheng,X.H., White,T.J., Shinsky,J.J., Adams,M.D. and Cargill,M.				
TITLE	Direct Submission				

844	QY	TCCTCAGCGCCAGCTGTGGGCGCGCTGGCAGAAACGAGCCGAGCTGCACCAACCCGG	903	1878	DB	CAGTCAGGACAAATGGGAGGATGTTGTGGTG	1931
807	DB	TGTTAAACAGCCGCTGTGGGCGAGGATATCAGAAACGCACAAAGAACCTGCACCAACCCAG	866	1930	QY	TTCTACTACTGCCAGCTGGAGGCCAGTGCCTGTCTTACCGAGAGCTGGGCGCT	1989
904	QY	GGCTCTCAACGGGGGCGCTTTCTGTGAGGGGAGAAATGTCATGACGGACCGTCTCCT	963	1935	DB	CCCTGTTACATTCAGCTGGATGCAGAGGCTTGCATATCTCTCACAGAGAACTCAGTACCT	1994
867	DB	CCCCACTCAATGGTGGGCGCTTCTGTGAGGGGACAGAGTGTGCAGAA	923	1990	QY	TTCCCTCTGGTGGAGAGGCCCTCAGCGTGGCTCCCGCAAGCGCTCAAGCTGCTTCTGT	2049
964	QY	CTCTCTTGTCTCTGTGAGCGGAGCTGGAGCCGCTGGAGCAAGTGTGCGCCGTGTGGG	1023	1995	DB	ATGCCCTGGTGGGCGAGTCCACCAACAGCAGCTGCAGCGCTCTTAACTGGCCATCT	2054
924	DB	CTACGTTATGTCCAGTGTAGTGGATGTTATGAGGCAATGTCTCAACCTGTGGGA	983	2050	QY	TTGGCCGGTGGCTGCACCTCCCTCGAGTACAAATCCGGGTCTACTGCTGTCATGACA	2109
1024	QY	TGGACTGCACCACTGGCGGAGCCGTGAGTGTCTCTGACCCAGCACCCCGCAAGCGGG	1083	2055	DB	TTGGGCCCTCTCTGCTCTCTTCCCTGGAGTACAGCAATTAGAGTCTACTGCTTGAATGACA	2114
984	DB	CTGATGACACCACTGGCGCAGGAGGAGTGTACAGCACACAGCCCCCAAGAAACGGGGA	1043	2110	QY	CCCAAGATGCACCTCAAGAGGCTGTGAGCTGGAGAGCAGCTGGGGGAGCAGCTGATCC	2169
1084	QY	AGGAGTGCAGGGGCACTGACTGTGAGACACCGGAACTGTACAGTGAACCTCTGTGTACACA	1143	2115	DB	CACAGATGCCCTGAGGAAGTTCTCAACTGGAGAGGCAATGGGAGGACAGCTCCTAG	2174
1044	DB	AGGACTGTGATGGCGCTGCTCTCCAAATCAAGAACTGCACATGAGTGGGTGTGCAATG	1103	2170	QY	AGGAGCCACGGGTCTGTGACCTCAAGAGCAGTTACACAACTGGCGCTATCCATCCAGC	2229
1144	QY	GTGCTTCTGGCCCTGAGACGCTGGCCCTCTATGTGGG	1200	2175	DB	AAGAAACCAAGGCTCTTCAATTTAAAGGCGAGCATCCAAACCTGGCGCTGTCTATTCTATG	2234
1104	DB	CTGCTCTGACTCAGATGATGTGGTCTCTACGTGGGAGTGTGATCGCTGTAAAGTCT	1163	2230	QY	ATGTGCCACAGCTCCCTGTGGAAGAGTAAGCTCCCTTGTGACCTACCAAGAGATCCCTTTT	2289
1201	QY	GCCTGGTCTGCTGTCTCTCATCTCTCGTTTATTGCGGAAAGAGGGGCTGG	1260	2235	DB	ACATCGGCCAATCCCTCTGGAAGAGCAAAATTTGCTGGCTAAGTATCAGGAATTTCCATTTT	2294
1164	DB	GTCTGGCGATCACTGTTGTGTGGCCCTGTTGTGTATCGGAAGAACCAACCGTGA	1223	2290	QY	ATCATCTCTGGAATGGCACGACGCGTACTTGTGCACTGCACCTTACCTCTGGAGCGTGTCA	2349
1261	QY	ACTCAGATGTGGTGACTGTGTCATCTTCACTCAGGCTTCCAGCCGTGAGCATCAAGC	1320	2295	DB	ACCACATCTGGAGTGGCTCTCAAGAAACCTCCACCTGCACCTTCACTCTGGAAGACTCA	2354
1224	DB	AGTCTGACATATTGACTCTCAGCACTCAATGGCGGCTTTTTCAGCCTGTGAACATCAAG	1283	2350	QY	GCCCCAGCACTAGTGAACCTGCCCTGTGAAGCTGTGGGTGTGGCAGTGGAGGGGCGACGGC	2409
1321	QY	CCAGAAAGCAGACAAACCCCATCTGCTCACATCCAGCCGAGCTCAGCACCAACCA	1380	2355	DB	GCCTAAACACAGTGGAACTGGTTTGCAAAACCTCTGTGTGGCGCAGGTTGAAGGAGAAAGGC	2414
1284	DB	CTGCCAGACAAAGATCTCTGGCTGTGCCCTCTGAC	1337	2410	QY	AGACCTTCAGCATCAACTTCAACATCACCAAGGACACAAAGTTTGTGAGCTGTGGCTC	2469
1381	QY	CTTACAGGGAGTCTCTGTCTCCCGGCGAGGATG	1435	2415	DB	AGATCTTTCAGCTCAACTGTACTGTGTGAGAGAACTTACTGGCATGCACTTACTCTCC	2474
1338	DB	TGTACAGGGAGCTCTATGCTCTGATGATGTCTCAGACAAATCCCAATGACCAACT	1397	2470	QY	TGGAGAGTGAAGCGGGGTCCAGCCCTGGTGGGCGCCAGTGCCTTCAAGATCCCTTCC	2539
1436	QY	CACCAATGGGCACTGTCTCAGCCCCCTGGGTGGCGGCGGCACACACTGTGCACACAGCTC	1495	2475	DB	TGGACCTGTAGTATACCATCACCACTGTACCAGCAACAGTGTCTTTCAGCATCTCTCTCC	2534
1398	DB	CTCCAATTTGAGACCACTACCCAACTTGAATAACAAAGTGTACAAACAGCTCAGTGTG	1457	2530	QY	TCATTCGGCAGAAAGATTAATTTCCAGCTGTGACCCACCCCTGTAGGCGGGGTGCCACTGGC	2589
1496	QY	TCCCACTC	1539	2535	DB	CTATCCGCGCAGAACTATGCAGCAGCTGGATGCTCCCTCAAAAGAGGGCCATGACTGGA	2594
1458	DB	TCACTCTCAGGATGACCTTGCAGGATCTCTATCCAAACTGTACCCAGATGACCCAGT	1517	2590	QY	GGACTTGGCCCGAGAACTCCACCTGGACAGCCATCTCAGCTTCTTTTGGCTCCCAAGCCCA	2649
1540	QY	-----AGAACTACTTCCGCTCCCTGCGCCCGAGGCA	1569	2595	DB	GGATGTGGGCCATTAACCTCAACCTGGACAGTACTTGAATTAATTTTGGCACCAATGGA	2654
1518	DB	CTTGTCTAGAGAAATGAGGCCCTTAACCTGAAGAACGAGAGCTCGCAAGACAGACTGACC	1577	2650	QY	GCCCCACAGCCATGATCTCAACCTGTGGGAGGCGGCGCACTTCCCAAGCGCAACCTCA	2709
1570	QY	CCAGCAACATGACTATGGGACTTCACTTCTCGGGGCGGCTGATGATCCCTAATA	1629	2655	DB	GCCCAACTGGCGTAACTCTGGATCTTTGGGAAGCACAGAACTTCCAGATGGAAACCTGA	2714
1578	DB	CATCTGTACAGCAATTTGGTACTTCACTCTCTTGGGGGTCACTCATCTATCTCTAAT	1637	2710	QY	GCCAGCTGGCTGTGACAGTGTGGACTGTGGCGCAGCCAGACGCTGGCCCTTCTTCAGTGT	2769
1630	QY	CAGGTATCAGCTCTCTATCCCGCAGATGCCATACCCGAGGAGATCTATGAGATCT	1689	2715	DB	GCATGCTGGCAGCCGCTCTCGAAGAAATGGGAAGACATGACAGTGTGTCTTGGCAG	2774
1638	DB	CAGGAGTAAGCTTGTGATTCCTCGCTGGGGCCATCTCCTCAGGGAGAGTCTATGAATG	1697	2770	QY	CGGAGCTGAGTGTGGA	2786
1690	QY	ACTCAGCTGCACAGCCGAGAGCTGAGGTTGCCCCCTAGCTGGCTGTGACAGCCCTGC	1749	2775	DB	CAGAAGGACAGTATTGA	2791
1698	DB	ATGTGACTACACAGMAAGAAATATGAGGCCCCCAATGGAAGACTCTCAGACCCCTAC	1757				
1750	QY	TGAGTCCCATCGTTAGCTGTGAGACCCCTGGCGTCTGTCTACCCGCGGCGAGTCACTCTGG	1809				
1758	DB	TTACCCCTGTGGTGTGAGCTGTGGGCGCTCTCTGAGCTCTGTGACCCGCGCTGTCTCCTCA	1817				
1810	QY	CTATGACCACTGTGGGGAGCCAGCCCTGTGACAGCTGTGAGCTGTGGCCCTCAAAAGACAGT	1869				

QY 670 ACCCAATGATATACATACAGCGGGAGACACAGCTGTGGTGGAGAGCGCGCTTGTCTG 729
Db 755 ATCGAACTTTTATATTAATCATCATCAACCTGATCATCAAGCAAGCCCGACTCTCAG 814
QY 730 ACAGCGCAACTACACCTCGTGGCCAAAGAACATCGTGGCAGCTCGCGCAGCGCTCGG 789
Db 815 ATACAGCAAAATTAACCTGTGTGGCCAAATATATTTGGCCAGAGAAAAGCACACAG 874
QY 790 CTGTGTCTCATCGTCTACGTGAACGCTGGGTGTGTGACGTGGACCGAGTGGTCCGCTGCA 849
Db 875 CCACGTGTATGTTAATGTGTGCTGTCCACCTGGACAGAGTGTGTGTGTA 934
QY 850 GGGCAGCTGTGGCGCGCTGGCAGAAACGAGCGGAGCTGCAACCAACCGCGCGCTC 909
Db 935 ACAGCGCTGTGGCGAGGATATCAAGAACGCAAGAACCTGCAACCAACCGAGCCCCAC 994
QY 910 TCAACGGGGCGCTTTCTGTGAGGGGAGAAATGTCATGACCGCACCGTCTCTCTGCG 969
Db 995 TCAATGGTGGGCGCTTCTGTGAGGGGAGAGTGTGAGAA---AATAGATGACTACGT 1051
QY 970 TTGTCTGTGTGACGGCAGCTGGAGCCCGCTGGAGCAAGTGGTGGCGCTGTGGGCTGGACT 1029
Db 1052 TATGTCCAGTGTGATGTAGTGGTCTTCAATGAGCAAAATGCTCAACCTGTGGGACTGAAT 1111
QY 1030 GCACCCACTGGCGGAGCGGTGAGTGTCTGACCCAGCACCCCGCAACGAGGGGAGGT 1089
Db 1112 GCACCCACTGGCGGAGGAGGTGTACAGCACGACCGCCCCAAGAACCGGGGTAAAGACT 1171
QY 1090 GCCAGGGCACTGACCTGGACACCCGCACTGTACAGTGAACCTCTGTGTACACAG----- 1144
Db 1172 GTATGGCGCTGTCTCTCAATCAAGAACTGCACTGATGGGCTGTGATGAGGAGTTCA 1231
QY 1145 -----TGCTTCTG 1152
Db 1232 TTTTACCCATTTCACTGAGCACAGACCCACAGAAATGAATATGAAATTTTCTTGTCTCTG 1291
QY 1153 GCCTGTGAGACGTGGCGCTCTATGTGG---CCTCATCGCGTGGCGCTGCTCGTCTCC 1209
Db 1292 ACTCAGATGATGGCTCTCTACGTGGGATGTGATCGCTGTAAACGCTGTCTGGGCA 1351
QY 1210 TGTGTCTGTGTGCTTCATCTCTGTTTATGTCGGGAAGAGGGGCTGGACTCAGATG 1269
Db 1352 TCACTGTGTGTGGCGCTGTGTGTATCGGAAGAACCCCGTGAATTTGAGTCTGACA 1411
QY 1270 TGGCTGACTGTGCTCACTTCACTCAGGTTCAGCCCGTCAAGATCAAGCCAGCAAG 1329
Db 1412 TCATTGACTCTCAGCACTCAATGGCGGCTTTCAGCCTGTGAACATCAAGGCTGCCAGAC 1471
QY 1330 CAGACAAACCCCATCTGCTCACCATCCAGCGGACCTCAGCACCAACCACTACCAGG 1389
Db 1472 AAGATCTCTGTGTGCTGCTCCCTGACC-----TCACCTCAGCTGCGAGCCATGTACAGG 1525
QY 1390 GCAGTCTGTGCTCCCGCAGGATG-----GGCCAGCCCAAGTTCAGCTCACCAGTGG 1444
Db 1526 GACCTGTCTATGCTCTGATGATGTCTCAGACAAATCCCAATGACCACTCTCCAATC 1585
QY 1445 GCACCTGTGACGCCCTCGGTGGCGCGCCACACATGTGACACAGCTCTCCACCTC 1504
Db 1586 TGGACCCACTACCCAACTTGAAATCAAAAGTGTACAAACAGCTCAGGTGTGTCACTCCTC 1645
QY 1505 -----TGAGGCCGAGGAGTGTCTCTCCGCTCTCCACCC----- 1539
Db 1646 AGGATGACCTTGGAGTGTCTCATCAAACTGTACCCCGAGTAGCCAGTCTTGTGTAG 1705
QY 1540 -----AGAACTACTTCCGCTCTCCCTCCCGGAGGACCAAGCAACA 1578
Db 1706 AGAATGAGGCCCTTAACCTGAAGAAACAGAGCTGCAAGACAGACTGACCCATCTGCA 1765
QY 1579 TGACCTTANGGACCTTCAACTTCTCGGGGGCGGTGATGATCCCTTAATACAGGTATCA 1638
Db 1766 CAGCAATTTGTACCTTCAACTCTCTTGGGGGTCACTCATCATTTCTTAATTCAGGAGTAA 1825
QY 1639 GCCTCTCATCCCCCGAGATGCCATACCCCGAGGGAAGATCTATGATCTACCTCAGCG 1698

Db 1826 GCTTGTCTGATTCGCGTGGGSCCATTCCTCAGGGAGAGTCTATGAATGTATGTGACTG 1885
QY 1699 TGCACAGCCCGAAGACGTGAGGTGGCCCTAGCTGGTGTGAGACCCCTGCTGAGTCCCA 1758
Db 1886 TACACAGAAAGAAAATATGAGGCCCCCATATGGAAGACTCTCAGACCCCTACTTACCCCTG 1945
QY 1759 TCGTGTAGCTGTGACCCCTCGCGTCTGCTCACCCGCCAGTCTATCTGGCTATGGACC 1818
Db 1946 TGGTGTAGCTGTGGGCTCTCTGGAGCTCTGTGACCCGCGCTGTATCTCTCACTCTGCATC 2005
QY 1819 ACTGTGGGAGCCAGCCCTGACAGCTGGAGCTGCGCTCAAAAGAGAGTGTGCGAGG 1878
Db 2006 ACTGTGCAGACCCAGCAGCAGGAGTGGAAATCCAGTCAAAAACAGGAGTGCAGG 2065
QY 1879 GCAGCTGGAGAGAGATGTGTGACCTGGGGGAGGAGCGCTCCACCTCTACTACT 1938
Db 2066 GACAATGGAGGATTTGTGGTG---GTGGGGAGGAGAACTTCAACCCCTCTTTACA 2122
QY 1939 GCCAGCTGGAGCCAGTGTCTGTACTGTCTTCAACGAGCAGCTGGGCGCTTTGGCCCTGG 1998
Db 2123 TTCAGCTGTGATGACAGGCTTGGCATATCTTACAGAGAACTCAGTACCTATGCCCCCTGG 2182
QY 1999 TGGGAGAGCCCTCAGCGTGGCTGCGCCAAAGCGCTCAAGCTGCTCTGTGTTGCGCCGG 2058
Db 2183 TTGGGCACTCCACCAACCAAGCAGCTGCCAAGCGTCTTAAACTGGCCATCTTTGGGCCCC 2242
QY 2059 TGGCCTGACACTCTCCCTCGAGTACAACTCCGGTCTACTGCTGCTGATGACACCCAGGATG 2118
Db 2243 TCTGTGTCTTCTTCCCTGGAGTACAGCTAGAGTCTACTGCTGCTGATGACACAGAGATG 2302
QY 2119 CACTCAAGGAGTGTGTGAGCTGGAGAGCAGCTGGGGGAGCAGCTGATCCAGGAGCCAC 2178
Db 2303 CCTGAAGGAGTTCTCAACTGGAGAGCAATGGAGAGCAGCTCTTAGAAGACCCA 2362
QY 2179 GGGTCTCTGCACTTCAAGGACAGTTACCAAACTGGCGCTTATCCATCCAGATGTCCTCA 2238
Db 2363 AGGCTCTTCTGTTTAAAGGAGCAGTCCCAAACTGGCGCTTCTATTATGACATCGCCC 2422
QY 2239 GCTCCCTGTGGAAGAGTGTGCTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 2298
Db 2423 ATTCCCTCTGGAAGAGCAAAATGCTGGTAAAGTATCAGGAAATTCATTTTACCACTCT 2482
QY 2299 GAAATGGCAGCAGCGGTACTTGCATGCACTTCACTTCACTTCACTTCACTTCACTTCACTT 2358
Db 2483 GAGTGGCTCTCAAGAAACCTCCACTGCACTTCACTTCACTTCACTTCACTTCACTTCACTT 2542
QY 2359 CTAGTACCTGCGCTGCAAGCTGTGGTGTGGAGGTGGAGGGCGACGGGCGAGGCTTCA 2418
Db 2543 CAGTGGAACTGGTTTGAACACTCTGTGTGGCGAGGTTGAAGAGAGAGGAGAGATCTTCC 2502
QY 2419 GCATCAACTTCAACATCAACAGGACACAAAGTTTGTGCTGAGCTGTGCTGCTGAGAGTG 2478
Db 2603 AGCTCAACTGTACTGTGTGAGAGAACCTTACTGGCATGACATTTACCTCTCTGAGACCTG 2662
QY 2479 AAGCGGGGTCCAGCGCTTGGGGCCCGAGTGGCTTCAAGATCCCTTCTCTCAITTCGGC 2538
Db 2663 CTAGTACCATCACCATGTGTACCGGACCAAGTGTCTTCAAGATTTCTCTCTCTCTCTCT 2722
QY 2539 AGAAGATAATTTCCAGCTGGACCCACCTGTAGGGGGGTGCGAGTGGGAGCTCTGG 2598
Db 2723 AGAAGCTATGACGACCTTGGATGGCCCTCAAAAGAGGAGGAGGAGGAGATGCTGG 2782
QY 2599 CCCAGAACTCCACCTGGAGACGATCTCAGCTTCTTGTCTTCAAGCCGAGCCCGACAG 2658
Db 2783 CCATTAACCTCAACCTGGAGAGTACTTGAATTTACTTTTGGCCCAATCAGAGCCCACTG 2842
QY 2659 CCATGATCTCAACCTGTGGAGGCGGAGCTTCCCAACCGCAACCTCAGGAGCTGG 2718
Db 2843 GCGTATCTTGGATCTTTGGAGAGCAGAACTTCCAGATGGAACCTGAGCATGTGG 2902
QY 2719 CTGAGAGTGTGGCTGGAGCTGGGCGAGCAGACCTGCGCTCTTTCACAGTGTGGAGGCTG 2778

Db	600	AAACAAGCCAGTGAACTGCACTCGAGAGCCTTCCTCTGCCACGACAGATCTACTTCAAGTGT	659
Qy	318	AAACGGGAGTGGGTGGCCACAGGTGGACCAAGTGAATCGAGCGCAGACACAGCGGAGCAGT	377
Db	660	AATGGCGAGTGGGTCAAGCAGAAATGACCAAGTCAACAGGAGAGCCTGGATGAGGCCACA	719
Qy	378	GGTGAGCGGACCATGGAGGTCCGCAATTAATGTCTCAAGCAGCAGAGTCCAGAGGTGTTC	437
Db	720	GGCTTGGGGTGGAGAGGTGCAGATCGAGGTGTCAAGCAGCAGCAAGTGGAGAACTCTTTC	779
Qy	438	GGGCTGAGGAATACTGGTCCAGTGGTGGCATGGAGTCTCTCGGCGCACCAAGAGT	497
Db	780	GGGCTCGAGGACTACTGGTGGCAGTGGTGGCTGGAGCTTTCGGGAACTACCAAGAGT	839
Qy	498	CAGAAGGCCTCATCCGCATAGCCAGATTCGCGAAGAACTTCGAGCAGAGCCGCTGGCC	557
Db	840	CGCGAGCCTACATCCGCATTGGCTACTTCGCAAGAACTTTCACGAGAGCCTCTGGGC	899
Qy	558	AAGGAGGTGTCCCTGGAGCAGGAGCATCGTGTGCTGCCCTCGCCGTCCACCGGAGGCACTCCCT	617
Db	900	AAGAGAGTACCTTGGATCATGAGTCTTCTGCAAGTGGCCGCCCAACGGAGGAGTGCT	959
Qy	618	CCAGCCGAGGTGGAGTGGCTCCGAAACGAGAACCTGGTGGACCCGCTCTCGACCCCAAT	677
Db	960	GTGGCTGAGGTGGAAATGGCTCAAGAAATGAAGATGTCAATTGACCCCGCTCAGGACACTAAC	1019
Qy	678	GTAATACATCAGCGGGAGCAGAGCCTGGTGGTGCAGAGCCGCGCTTGTCTGACACGGCC	737
Db	1020	TTCCGTGCTCACTATGACCAAACTCATCATTCGCCAGCGGCGCTCTCAGACAGGCC	1079
Qy	738	AACTACACCTCGTGGCCAAAGAACTCGTGACAGTCTGCCGACGCGCTCCGTGTGTGTC	797
Db	1080	AACTACACCTGTGGCCAAAGAAATATCGTGGCCAAAGCGCCGAGCAGCGCGGCCACAGTC	1139
Qy	798	ATCGTCTACGTGAACGTTGGTGGTGTGCACTGGACCGAGTGTCCGCTCTCAGCGCCAGC	857
Db	1140	ATCGTCTATGTGAATGGAGGCTGGTGCAGCTGGGACAGTGGTCAACCTGTTTCAATCGC	1199
Qy	858	TGTGGCGCGCTGGCAAAACGGACCGGAGCTGACCAACCCGGCGCCCTCTCAAGCGG	917
Db	1200	TGTGGCCGAGCTGGCAGAACGCTACTCGGACCTGCACCAATCAGCCCCACTCAATGGA	1259
Qy	918	GGCGTTTCTGTGAGGGCAGAAATGTCCATGACCGCACCGTCTCTCTGCTGTGTCTCT	977
Db	1260	GGCGCTTCTGTGAGGGACAGGCCCTTCCAGAAGACAGCTTGACACCGGTGTG---CCCA	1316
Qy	978	GTGGACGACAGCTGAGAGCCCTGGAGCAAGTGTGTGGCTGTGGCTGGACTGCACCCAC	1037
Db	1317	GTGGATGAGCGTGCACCGATGGAGCAAGTGGTCTGCCTGCACACAGAGTGTGCGCAC	1376
Qy	1038	TGGCGGAGCCGTGAGTGTCTGACCCAGCAACCCGCAACGAGGGGAGGAGTGCAGGCC	1097
Db	1377	TGGCGCAGCCGAGTGAATGSCACCGCACCCCAAGACCGGAGCCGTGATGTCAGCGGG	1436
Qy	1098	ACTGACCTTGGACACCCGCAACTGTACACAGTCACTCTGTGTACACAGTCTTCTGGCC--	1155
Db	1437	ACGCTACTTGATCCAGAAGCTGCATGATGGCTGTGCGTGTGTAATCAGAGAACTCTA	1496
Qy	1156	-----CTGAGGACGTGGGCCCTCTATGTGGGCCTC	1184
Db	1497	AAAGCACCCTAAAGGCCACCCCTGGAGACATCGGAGATGTGGCAGTGTACCGAGGCCTT	1556
Qy	1185	ATCGCGGTGGCGCTCTGCTGGTCTGTGCTGTCTGTGC---CTCATCTCGTTTATTGC	1241
Db	1557	GTGGTGGCGCTTTTGTGGTGTAGCGTCTCTATGGCCGAGGAGTGAATGTAATACGG	1616
Qy	1242	CGGAGAAGAGGGGCTGGACTCAGATGTGGCTGACTCGT---CCATTTCTCACTCAGGC	1298
Db	1617	AGAACTGCCGGACTTCGACACGGACATCACCGACTCTCTGCGGCCCTCACTGGTGGC	1676
Qy	1299	TTCCAGCCCGTCAAGATCAAGCCGAGAAAGCAGACAACCCCATCTGCT-----CACC	1352
Db	1677	TTCCACCCCTGCAACTTCAAGATCTCAAGCTCAAGGCCCAACACCCGAGCTCTCACCGTCC	1736

QY	1353	ATCCAGCCGCACTCAGACCAACCACTACAGGCGAGTCTCTGTCCCGCGAGGAT	1411
Db	1737	GCOCCTCCAGACTAACGGCCAGTGTGGGATCTACCGCGGGCTGTGTATGCCCTGCAG	1796
QY	1413	GGGCCAGCCCCAAGTTCCAGCTCACAAATGGGCACCTGCT	1453
Db	1797	GACTCGCCGCAAGATCCCATGACTAATTGCCCCCTGCTGATCCCTGCCAGCCTC	1856
QY	1454	-----C	1454
Db	1857	AAGATCAAGGTTCTAATCTCCAGCACCATCGGTTCTGGGTCTGGCTGCTGATGAGCC	1916
QY	1455	AGCCCCCTGGTGGCGGCCACACACTGCACACAGACTCTCCCACTCTCAGGCCGAG	1514
Db	1917	GACCTGCTGGGTGTCCTCCCGCGGCAGTACCCAGGCGATTCTCCCGGGACACCCAT	1976
QY	1515	GAGTTGCTCCCGCCTCTCCACCCAG-----AACTACTTTCGCTCCCTGCCCCGAGGC	1568
Db	1977	TTCTCGACCTGGCAGTGCAGCCTTGGTTCCAGCACCTCTCGGCCCTACTCTGGGAC	2036
QY	1569	ACAGCAACATGACCTATGGAGCTTCAACTCTCTGGGGCGCGTGATGATCCCTAAT	1628
Db	2037	CCAGCAGCAGTGTACGCGGACCTTGGTTGGCTGGGAGAAAGCTGAGCCTCCCGGC	2096
QY	1629	ACAGTATCAGCCTCCTCATCCCCCAGATGCCATACCCGAGGGAGATCTATGATC	1688
Db	2097	ACAGGGTTCAGCCTGTTGGTACCAATGGAGCCATCCCCAGGGCAAGTTCTATGACCTG	2156
QY	1689	TACCTCAGCTGCAACAAGCCGGAAGACGTGAGTTGCCCTAGCTGGCTGTGAGACCTTG	1748
Db	2157	TATCTACATATCAACAAGGCGGAAGCACCTCCCACTTTCAGAAGGTTCCCAAGACAGTA	2216
QY	1749	CTGAGTCCCATCGTTAGTGTGGACCCCTGGGTCCTGCTCAACCCGCGCAGTCACTCGT	1808
Db	2217	TTGAGCCCCCTCGGTGACCTGTGGGCCCAAGGCTACTCTCTGTGCGCCCTGTGCTCTC	2276
QY	1809	GCTATGGACCACTGTGGGGACCCAGCCCTGCACAGCTGGAGCCTTGCGCCTCAAAAGCAG	1868
Db	2277	ACCGTGCCCCACTGTGTGAAGTCATCGCTGGAGACTGGATCTTTCACTCAAGACCCAG	2336
QY	1869	TGTCGAGGCGAGCTGGAGCAGGATGTCTGCACCTGGCGGAGGAGCGCCCTCCAC	1928
Db	2337	GCCCATCAGGCCACTGGG---AGAGGTGGTGACTTGGATGAGGAGACCTCAACACA	2393
QY	1929	CTCTACTACTCCAGCTGGAGGCCAGTCCCTGTACGTTCTTACCGAGCAGCTGGGCCG	1988
Db	2394	CCCTGCTACTGCAGCTGGAGGCTAAGTCTCTGCCACATCTCTGTGGACCACTGGTTC	2453
QY	1989	TTTGCCCTGTGGAGAGGCCCTCAGCGTGGCTCGGCCAAGCCCTCAAGCTGCTTCTG	2048
Db	2454	TACGTATTTCATGGCGAGTCTCTCTGCTCTGAGTCAAGCGGCTCCAGCTGGCCATC	2513
QY	2049	TTTGGCGCGGTGGCTGCACCTCCTCGAGTACAACATCCGGGTCTACTGCTCATGAC	2108
Db	2514	TTGCCCCAGCCCTCTGCACCTCTCTGGAGTATAGCCTCAGGCTCTACTGTCTGGAGAC	2573
QY	2109	ACCCACGATGCATCAAGGAGGTGGTGCAGCTGGAGAGCAGCTGGGGGACAGCTGATC	2168
Db	2574	ACACCTGTAGCACTGAAGGAGGTCTCTGGAGCTGGAGAGGACTCTGGGTGGCTACTTGGT	2633
QY	2169	CAGGAGCCACGGGTCTCTGACTTCAAGAGACAGTTACCACACCTCCGCTATCCATCCAC	2228
Db	2634	GAGGAGCCCAAGCCTTTGCTCTTTAAGAGCAGTTACCAACCTTACGCTCTCCCTCCAT	2693
QY	2229	GATGTGCCAGCTCCCTGTGGAAGTAGTCCCTTGTCACTACCAAGAGATCCCCCTT	2288
Db	2694	GACATCCCCCATGCCCCACTGGAGGAGCAACTCTGGCCAAAGTACCAGAGATTCCCTTC	2753
QY	2289	TATCATCTGGAATGGCAGCGAGCGGTACTTGCNCTGCACCTTCAACCTCGAGCGGTGC	2348
Db	2754	TACCAAGTCTGGAATGGCAGCAGAGCCCTGCACTGCACTTTCACCTCAGAGGAT	2813

Db 1030 ACAGCAGCGCGGATCTACCGGAGACCCGTGTATGCCCCTGCAGGACTCCACGACAAA 1089
Qy 1428 TTCAGCTCACAATGGCACTGCTCAGCCCC----- 1460
Db 1090 ATCCCCATGACCAACTCTCTCTGCTGTGACCCCTTACCAGCCCTTAAGGTCAAGGTCTAC 1149
Qy 1461 -----CTGGGTGGC 1469
Db 1150 AGCTCCAGCACACCGGCTCTGGGCCAGCCCTGGCAGATGGGGCTGACCTGTCTGGGGTC 1209
Qy 1470 GCGCGCACACACTGCACAGACTCTCCACCTCTGAGGCGGAGAGTGTCTCTCCGC 1529
Db 1210 TTGCGCTGGCAGATACCTAGCAATTTGCGCGGAGACACCCACTTCTGCACTGCGC 1269
Qy 1530 CTCTCCACCCAGAACTAC-----TTCCGCTCCCTGCCCGGAGGACCCAGCAACATGACC 1583
Db 1270 AGCGCAGCTCGGTTCCAGCAGCTCTTGGGCTTGCCCGAGACCCAGGAGCAGCGTC 1329
Qy 1584 TATGGGACCTTCAACTTCTCTGGGGCGGCTGATGATCCCTTAATACAGGTATGACCTC 1643
Db 1330 AGCGCAGCTTTGGCTGCTGGTGGGAGGCTCAGCATCCCCGGCACAGGGGTGAGCTTG 1389
Qy 1644 CTATCCCCCAGATGCCATACCCCGAGGAGAGATCTATGATCTTACCTCAGCTGCAC 1703
Db 1390 CTGGTGCCCAATGAGAGCAATCCCCAGGGCAAGTTCTACGAGATGATCTACTCATCAC 1449
Qy 1704 AAGCGGAGAGCTGAGGTGTCCTCCCTAGCTGTGCTGCTCAGACCTCTGAGTCCCATCGTT 1763
Db 1450 AAGCGAGAAAGTACCCTCCCGCTTTTCAAGAGGAGCCAGACAGATTTAGAGCCCTCGGTG 1509
Qy 1764 AGCTGAGACCCCTGGGCTCTGCTCACCAGGACAGTCACTCTGGCTATGACCACTGT 1823
Db 1510 ACCTGTGACCCACAGGCTCTGCTGTGCGCCCGCTCATCTCACCATGCCCCACTGT 1569
Qy 1824 GGGAGGCCAGCCCTGACAGCTGGAGCTGCGCTCAAAAAGCAGTGGCGAGGCGACG 1883
Db 1570 GCGAGTCAGTGCGCGGTGACTGATCTTTAGCTCAAGCCAGGCCACACAGGCGCAC 1629
Qy 1884 TGGGAGCAGGATGTGTGACCTGGCGGAGGCGCCCTCCACCTCTACTACTGCGCAG 1943
Db 1630 TGGG---AGGAGTGTGTGACCTGGATGAGGAGACCTCAACACACACCTGTACTGCCAG 1686
Qy 1944 CTGGAGGCGAGTCTGCTACCTCTTACCGAGCAGCTGGCGCTTTCCTCTGTTGCGGA 2003
Db 1687 CTGGAGCCAGGCGCTGTACATCTCTGTGGACCACTGGGACCTTACTGTGTACGGGC 1746
Qy 2004 GAGGCTCAGGCTGCTGCCGCCAAGCGCCTCAAGCTGTCTGTGCGCGGTGGCC 2063
Db 1747 GAGTCTATTCGGCTCAGCAGTCAAGCGCTCCAGCTGGCGCTTTCGCCCGCGCTC 1806
Qy 2064 TGCACCTCCCTCGAGTACAACTACCGGCTCTACTGCTGATGACACCCAGGATGACTC 2123
Db 1807 TGCACCTCCCTCGAGTACAGCTCCGGCTCTACTGCTGGAGGACACGCTGTAGCACTG 1866
Qy 2124 AAGGAGTGTGAGTGTGAGAGCAGCTGGGGGACAGCTGATCCAGGACCAAGGCTC 2183
Db 1867 AAGGAGTGTGAGTGTGAGGAGTCTGGGCGGATCTTGTGAGGAGCGGAAACCG 1926
Qy 2184 CTGCACCTTCAAGGACAGTTTACCAACCTCGCGCTATCCATCCAGCATGTGCCAGCTCC 2243
Db 1927 CTAAATTTCAAGGACAGTTTACCAACCTCGCGCTCTCCCTCATGACCTCCCGATGCC 1986
Qy 2244 CTGTGAAAGTAACTCTCTGTGACTACAGGAGATCCCTTTTATCATCTGGAAT 2303
Db 1987 CATTTGAGGAGCAAGCTGTGCGCAATACCAAGAGATCCCTTCTATCATCTTTGAGT 2046
Qy 2304 GGCACGCGGTACTTGTGACTGACCTTCACTCTGAGCTGTGAGCGCCAGCACTAGT 2363
Db 2047 GGCAGCGCAAGGCCCTTCACTGCACTTTTCACTCTGAGGAGGACAGCTTGGCTCCACA 2106
Qy 2364 GACCTGGCTGCAAGCTGTGGGTGTGGCAGGTGGAGGCGACGGGAGAGCTTCAGCATC 2423

Db 2107 GAGCTCACCTCAAGATCTGGTGGCGCAAGTGAAGGGAGGGCCAGATATTCCAGCTG 2166
Qy 2424 AACTTCAACATCAACAAG---GACACAAGTTTGTGTAGCTGTGGCTCTGGAGAGTGAA 2480
Db 2167 CATACCACTCTGGCAGAGACACCTGCTGGCTCCCTGGGACACTCTCTGTCTGCCCCCTGGC 2226
Qy 2481 GCGGGGTCCAGAGCCCTGGTGGGCCCCAGTGCCTTCAAGATCCCCCTTCTCATTTGGCAG 2540
Db 2227 AGCACTGTCAACACCCAGCTGGGACCTTATGCTTCAAGATCCCACTGTCATCCGCCAG 2286
Qy 2541 AAGATAATTTCCAGCTGGACCCACCTGTAGGCGGGGTGGCGACTGGCGGACTCTGGCC 2600
Db 2287 AAGATATGACACAGCCTAGATGCCCACTCAACCGGCAATGACTGGCGGATGTAGCA 2346
Qy 2601 CAGAACTCCAGCTGGACAGCCTCTCAGCTTCTTTGGCTCCAGCCAGCCCCCAGCC 2660
Db 2347 CAGAGCTCTCTATGGACCGGTACCTGAATTCTTTGCCACCAAGCAGGCCCCAGGTT 2406
Qy 2661 ATGATCTCTCAACCTGTGGGAGCGGCGACCTTCCCACAGCGCAACCTCAGCCAGCTGGCT 2720
Db 2407 GTGATCTGGACCTCTGGGAGCTCTGACGAGGAGCTGGGAGCTCAACAGCTGGCG 2466
Qy 2721 GCAGCAGTGGTGGACTGGGCGAGCCAGCAGCTGGCCTTTTCAAGTGTGGAGGCTGAG 2780
Db 2467 AGTGCCTTGGAGGAGATGGCAAGAGTGAATGCTGTGGTGTGGCCACCGACGGGAC 2526
Qy 2781 TGCTGA 2786
Db 2527 TGCTGA 2532

RESULT 6

AY406492

LOCUS

DEFINITION

Pan troglodytes UNCS gene, VIRTUAL TRANSCRIPT, partial sequence,

genomic survey sequence.

ACCESSION

AY406492

VERSION

AY406492.1 GI:39762466

KEYWORDS

GSS.

SOURCE

Pan troglodytes (chimpanzee)

ORGANISM

Pan troglodytes

REFERENCE

1 (bases 1 to 2802)

AUTHORS

Clark, A.G., Glanowski, S., Nielson, R., Thomas, P., Kejarawal, A.,

Todd, M.A., Tanenbaum, D.M., Civallo, D.R., Lu, F., Murphy, B.,

Perriera, S., Wang, G., Zheng, X.H., White, T.J., Shinsky, J.J.,

Adams, M.D. and Cargill, M.

TITLE

Inferring nonneutral evolution from human-chimp-mouse orthologous

gene trios

JOURNAL

Science 302 (5652), 1960-1963 (2003)

PUBMED

14671302

REFERENCE

2 (bases 1 to 2802)

AUTHORS

Clark, A.G., Glanowski, S., Nielson, R., Thomas, P., Kejarawal, A.,

Todd, M.A., Tanenbaum, D.M., Civallo, D.R., Lu, F., Murphy, B.,

Perriera, S., Wang, G., Zheng, X.H., White, T.J., Shinsky, J.J.,

Adams, M.D. and Cargill, M.

TITLE

Direct Submission

JOURNAL

Submitted (16-NOV-2003) Celera Genomics, 45 West Gude Drive,

Rockville, MD 20850, USA

COMMENT

This sequence was made by sequencing genomic exons and ordering

them based on alignment.

FEATURES

Location/Qualifiers

source

1..2802

/organism="Pan troglodytes"

/mol_type="genomic DNA"

/db_xref="taxon:9598"

<1..>2802

/gene="UNC5C"

/locus_tag="HCM2575"

ORIGIN

Query Match 26.5% Score 763.2; DB 29; Length 2802;

Best Local Similarity 52.9%; Pred. No. 8e-118; Matches 1456; Conservative 0; Mismatches 1218; Indels 78; Gaps 7;			
QY	98	GCCCCGCGCTGTGGCCAGCGCTCCTGGGCATAGTCTCGCGCTTGGCTCGCGGCTCGGG	157
Db	66	GCTCGTGCTACTCGTCCCTGGCCCTGCTCAGCGCCAGCGCACTGGCTCGCGCCCAAGA	125
QY	158	TGCCACGACAGTGCACCGTGGCCAAACCCAGTGCCTGTGCGCAACCCCGACCTGCTTCC	217
Db	126	TGATGACTTTTTCATGAATCCCAAGAACTTTTCTTCGGATCCACCTGAGCCTCTGCC	185
QY	218	CCACTTCTCTGTGGAGCCCGAGGATGTGTACATCGTCAAGAAACCAAGCCAGTGTCTGT	277
Db	186	ACATTTCCTTATTGAACCTGAGGAAGCTTATTGTGAAGATAAGCCCTGTGAACCTGTA	245
QY	278	GTGCAAGGCC-----GTGCGCCGACGACGATCTTCTTCAAGTGCAACGGGGAGTGGT	331
Db	246	CTGTAAAGCAAGCCCTGNN	305
QY	332	GCGCAGGTGGACCACTGTATCGAGCGCAGCAGACAGCGGAGCAGTGTGTGAGCGCACCAT	391
Db	306	NN	365
QY	392	GGAGTCCGCAATTAATGTCTCAAGCGCAGAGTGCAGAAAGTGTTCGGGCTGGAGGAATA	451
Db	366	NN	425
QY	452	CTGGTGCCAGTCCGTGGCATGGAGCTCTCTCGGCGACCAAGAGTCAGAAAGCCTACAT	511
Db	426	NN	485
QY	512	CCGCATAGCCAGATTGCGCAAGAACTTCGAGCAGAGCCGCTGGCCAAAGAGTGTCCCT	571
Db	486	NN	545
QY	572	GGAGCAGGCATCGTGTGCTGCTGCGCTCCACCGGAGGCACTCCCTCCAGCGGAGTGA	631
Db	546	GGAACAGGAAGTCTTACTCCAGTGTGCGCCACTGAAAGGGATCCCAAGTGGCTGAGNNNN	605
QY	632	GTGGCTCCGGAACGAGGACCTGTGTGAGCCGCTCCCTGGACCCCAATGTATACATCAGCG	691
Db	606	NNNNNNGAAAAATGAAGACATAATTGATCCCGTTGAAGATCGGAATTTTATTACTAT	665
QY	692	GGACACAGCCTGTGTGTGCGCAGCGCCGCTTGTGACACGCGCCAACTACACCTGCGT	751
Db	666	TGATCACAACTCATATAAGCAGGCGCCGACTCTCTGATCTGCAATTTATACCTGTGT	725
QY	752	GGCCAAAGAACATCGTGGCAGCGTGGCGCAGCGCTCCGCTGCTGTCTATCTTACGTGA	811
Db	726	TGCCAAAAACATTTGTGCCAAGAGGAAGTACAACTGCCACTGTCTATGTCTATGTCAA	785
QY	812	CGGTGGGTGTGCACGTGGACCGAGTGGTCCGCTCTGACGCGCCAGCTGTGGCGCGGCTG	871
Db	786	CGGTGGCTGTGCCACTGGACGAGTGGTCTGTGTGTAACTGCTGTGGAAGAGGTA	845
QY	872	GCAGAAACGGAGCCGAGCTGCACCAACCGCGGCTCTCAACGGGGGCGCTTCTGTGA	931
Db	846	TCAGAAACGTACAAAGACTTGTACCAACCGCGCACTCAATGGGGGTGCTTCTGTGA	905
QY	932	GGGGCAGAAATGCCATGACCGCAGCTCTCTCTCTGTGTGTCTGTGTGTGTGTGTGTGT	991
Db	906	AGGCGAGTGTGCAGAA---AATAGCCTGTACTACGTTATGCCAGTGGATGGCAGGTG	962
QY	992	GAGCCGTGGAGCAAGTGTGGCCCTGTGGCTGGACTGCACCACTGGCGGAGCGGTGA	1051
Db	963	GAGCCATGGAGCAAGTGTGTCTATTGTGGAACAGTGCACCCACTGGCGCAGGAGGA	1022
QY	1052	GTGCTCTGACCCAGCACCCCGAACCGAGGAGGAGTGCAGGGCACTGACTGGACAC	1111
Db	1023	GTGCACGGCGCCAGCCCAAGAAATGGAGCAAGGACTGCGATGGCTGTCTTGTCAATC	1082
QY	1112	CCGCAACTGTACAGTGAACCTCTGTGTACACAGTGTCTTGTGGCCTGTGAGGACGTGGCCT	1171
Db			
1083	CAMAGAACTGCACGTATGGGCTTTGCAATGCAGACTGCTCTGTGATTCAGATGATGTGCTCT	1142	
QY	1172	CTATGTGGG---CCTCATCGCGTGGCGCTCTGCTGCTGCTGCTGCTGCTGCTCAT	1228
Db	1143	CTATGTGGGATTTGTATAGCAGTGTGTTGCTGGCGATCTCTGTAGTGTGGCCTT	1202
QY	1229	CTCTGTTTATTGCCGGAAGAAGAGGGGCTGGACTCAGATGTGGCTGACTGCTGCATTCT	1288
Db	1203	GTTTGTGTATCGGAAGAATCATGTAATTTGAGTCAGATATTTATGACTCTTTCGGCACT	1262
QY	1289	CACCTCAGGCTTCCAGCGCTCAGCATCAAGCCCAAGCAGACAAACCCCATCTGCT	1348
Db	1263	CAATGGGGCTTTTCAGCCTGTGAACATCAAG-----GCAGCAAGACAAGATCTGCT	1313
QY	1349	CACCATCCAGCCGAGCTCAGCACCAACCACTACCAAGGCGAGTCTCTGTCCCCGGCA	1408
Db	1314	GGCTGTATACCCCGAGACTCAGCTCAGCTCAGCCATGTACAGAGGACTGTCTATGCCCT	1373
QY	1409	GGATGGCCCGAGCCCAAGTTCCAGCTCACCAATGGGCACTGCTCAGCCCCCTGGGTGG	1468
Db	1374	GCATGACGTCTCAGACAAATCCCAATGACCAACTCTCCAATTTCTGGATCCACTGCCCAA	1433
QY	1469	CGCGCCGACACACTGCACACAGCT-----CTCCCACTCTCAGGCCGAGGA	1516
Db	1434	CCTGAAATCAAGTGTACAACTCAGGTGCTGTACCCCCCAAGATGACCTCTCTGA	1493
QY	1517	GTTCGTCTCCGCTCTCCACCCAGAACTACTTCCGCTCCCTGCCCGAGGACCAGCAA	1576
Db	1494	GTTTACGTCCAAGCTGTCCCTCAGATGACCCAGTCACTGTTGGAGATGAAGCCTCAG	1553
QY	1577	CATGA-----CTATGGGACCTT	1594
Db	1554	CCTGAAGAACACAGACTGTAGCAAGGCAGACTGATCCATCTGTACCGCATTTGGCAGCTT	1613
QY	1595	CAACTTCTCGGGGCGGCTGATGATCCCTAATACAGGTATCAGCTCTCTCATCCCCC	1654
Db	1614	CNNNTGCTGGNNNNACCTTATTTGTTCCCAATTCAGAGTCACTGCTGTGATTTCCGCG	1673
QY	1655	AGATGCCATACCCGAGGGAAGATCTATGAGTCTTACCTCAGCTGCAAGCCGGAAGA	1714
Db	1674	TGGGCCCATTTCCCAAGGAGAGTCTACGAAATGTATGTACTGTACACAGGAAGAAAC	1733
QY	1715	CGTGAAGTTGCCCTAGCTGCTCAGACCTGCTGAGTCCCATGCTGTAGCTGTGGACC	1774
Db	1734	TATGAGGCCACCCATGGATGACTCTCAGACACTTTTGACCCCTGTGTGAGTGTGGGCC	1793
QY	1775	CCCTGGCGTCTGCTCACCCGCGCAGTCACTCTGGCTATGACCACTGTGGGAGCCGAG	1834
Db	1794	CCGAGGACTGTCTCACCCGCGCTGCTCTTACTATGATCACTCTGCGCAGACCCCAA	1853
QY	1835	CCCTGACAGCTGGAGCCTTGGCCCTCAAAAGCAGTCTGTGCGAGGCGAGCTGGAGCAGGA	1894
Db	1854	TACCGAGGACTGGAAAAATCTGCTCAAGAACCCAGGACGACAGGACAGTGGG---AGGA	1910
QY	1895	TGTGCTGCACCTGGCGAGGAGGCGCCCTCCCACTCTACTCTACTCTGCCAGCTGGAGGCCAG	1954
Db	1911	TGTGGTGGTGGTGGGGAGGAAAACTTACACCCCTGCTTACATTCAGCTGGATGCAGA	1970
QY	1955	TGCCTGTACTGCTTTCAACGAGCAGTGGCGCTTTTGGCTGTGGGAGGAGGCCCTCAG	2014
Db	1971	GGCTGGCCACATCTCTCAGAGAACCTCAGACCTTACGCCCTGGTAGGACATTCACACAC	2030
QY	2015	CGTGGCTGCCCGAAGCGCTCAAGCTGCTTGTGTTGGCGGCTGGCTGCACTCCCT	2074
Db	2031	CAAAAGCGCTGCGAAGCGCTCAAGCTGGCCATCTTTTGGGCCCCCTGTGCTGCTCTCGCT	2090
QY	2075	CGAGTCAACATCCGGGTCTACTGCTGCATGACACCCAGATGCACTCAAGGAGGTGGT	2134
Db	2091	GGAGTACAGCATCCGAGTCTACTGTCTGGATGACACCCAGGATGCCCTGAAGGAAATTTT	2150
QY	2135	GCAGCTGGAGAGCAGCTGGGGGACAGTGTATCCAGGAGCCACGGGCTCTGCACTTCAA	2194
Db	2151	ACATCTTGAGAGACAGATGGGAGGACAGCTCTTAGAAGAACCTTAAGGCTCTTCAATTTAA	2210

```
QY 2195 GGACAGTTACCAACACCTGCGCCCTATCATCCAGATGTCGCCAGCTCCCTGTGGAGAG 2254
Db 2211 AGGCAGCACCCACACCTGCGCCCTGTCAATTCAAGATATCGCCCAATCCCTCTGGAAGAG 2270
QY 2255 TAAGTCCTCTGTGACGTACAGGAGATCCCTTTTATCATCTGGAATGGCAGCAGCG 2314
Db 2271 CAATTTGCTGGCTAAATATCAGNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNN 2330
QY 2315 GTACTTGCACATGACCTTACCCCTGGAGCGTGTGACGCCCGACACTAGTAGACCTGGCGCTG 2374
Db 2331 NAACCTGCACTGACCTTCACTCTGGAAGATTTAGCCTGGAACAGTGGAGCTGGTTG 2390
QY 2375 CAAGCTGGGTGTGGAGTGGAGGCGAGCGGAGAGCTTTCAGATCAACTCAATCAAT 2434
Db 2391 CAAACTCTGTGTGCGGCGAGGTGGAAGGAGAGGCGAGATCTTCCAGCTCAACTGCACCGT 2450
QY 2435 CACCAAGGACACAAAGTTTGTGCTGCTGCTGAGAGTGAAGCGGGGTCCACG 2494
Db 2451 GTCAGAGAACTACTTGGCATCGATTGCGCGTGTGGATCTTGGCAACACCATCACAC 2510
QY 2495 COTGTGGGCCCCAGTGCCTTCAAGATCCCTTCTCATTTGGCGAGAGATAATTTCCAG 2554
Db 2511 GGTACGCGGCCCGAGTCTTTCAGCATCCCTCTCCCTATCCGCGCAGAAAGCTCTGTAGCAG 2570
QY 2555 CCGTGAACCCACCTGTAGCGGGGTGCCGACTGGCGGACTCTGGCCCGAGAACTCCACCT 2614
Db 2571 COTGTATGCCCCCGACAGAGAGGCCATGACTGGAGGATGCTGGGCCATAAGCANNNNNN 2630
QY 2615 GGACAGCCATCTCAGCTCTTTGCTCCAAAGCCCGCCAGCCACAGCCATGCTCAACCT 2674
Db 2631 NNNNNNGTCTTGAATATTCTTGCCACCAATCCAGCCCACTGGGCTGATCT 2690
QY 2675 GTGGAGGCGCGCACTTCCCAAGCGCAACTCAGCCAGCTGGCTGCGAGAGTGGCTGG 2734
Db 2691 TTGGGAAGCACAGAACTTCCAGATGGAACCTGAGCATGCTGGCAGCTGTCTTGAAGA 2750
QY 2735 ACTGGCCAGCAGAGCTGGCTCTTACAGTGTGGAGGCTGAGTGTGA 2786
Db 2751 AATGGGAAGACATGAACCGTGGTGTCTTAGCAGCAGAGGCGAGATTAA 2802
```

```
RESULT 7
BI758231 1034 bp mRNA linear EST 25-SEP-2001
LOCUS 603029876F1 NIH_MGC_114 Homo sapiens cDNA clone IMAGE:5200171 5',
DEFINITION mRNA sequence.
ACCESSION BI758231
VERSION BI758231.1 GI:15749809
KEYWORDS EST.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
1 (bases 1 to 1034)
NIH-MGC http://mgi.nci.nih.gov/
National Institutes of Health, Mammalian Gene Collection (MGC)
Unpublished (1999)
Contact: Robert Strausberg, Ph.D.
Email: cgabbs-r@mail.nih.gov
Tissue Procurement: Life Technologies, Inc.
CDNA Library Preparation: Life Technologies, Inc.
CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
DNA Sequencing by: Incyte Genomics, Inc.
Clone Distribution: MGC clone distribution information can be
found through the I.M.A.G.E. Consortium/LLNL at:
http://image.llnl.gov
Plate: LLAM1501 row: g column: 20
High quality sequence stop: 793.
Location/Qualifiers
1..1034
/organism="Homo sapiens"
/mol_type="mRNA"
```

FEATURES

source

```
/db xref="taxon:9606"
/clone="IMAGE:5200171"
/lab host="DH10B"
/clone_lib="NIH_MGC_114"
/note="Organ: brain; Vector: pCMV-SPORT6; Site_1: NotI;
Site_2: EcoRV (destroyed); RNA source anonymous pool of 6
male brains, age range 23-27 yo. Library is oligo-dr
primed and directionally cloned (EcoRV site is destroyed
upon cloning). Average insert size 1.5 kb, insert size
range 1-3 kb. Library is normalized and enriched for
full-length clones and was constructed by C. Gruber
(Invitrogen). Research Genetics tracking code 019. Note:
this is a NIH_MGC Library."
```

ORIGIN

```
Query Match 26.2%; Score 756.2; DB 12; Length 1034;
Best Local Similarity 89.7%; Pred.No. 7.5e-117; Indels 16; Gaps 8;
Matches 905; Conservative 0; Mismatches 88;

QY 191 GCCTGCTGCCAACCCGACCTGCTTCCCCACTTCTGTTGGAGCCGAGGATGTGTACAT 250
Db 1 GCCTGCTGCCAACCCGACCTGCTTCCCCACTTCTGTTGGAGCCGAGGATGTGTACAT 60
QY 251 GGTCAAGAAACAGCCAGTGTCTTGTGTGCAAGCCGCTGCCGCCAGCATCTTCTT 310
Db 61 GGTCAAGAAACAGCCAGTGTCTTGTGTGCAAGCCGCTGCCGCCAGCATCTTCTT 120
QY 311 CAAAGTGCACCGGGAGTGGTGGCCAGTGCACCACTGATCGAGCGCAGCAGACAGCG 370
Db 121 CAAAGTGCACCGGGAGTGGTGGCCAGTGCACCACTGATCGAGCGCAGCAGACAGCG 180
QY 371 GAGCAGTGTGAGCCGACCATGAGGTCCGCAATTAATGTCTCAAGGACAGAGTGCAGAA 430
Db 181 GAGCAGTGGGCTGCCACCATGAGGTCCGCAATTAATGTCTCAAGGACAGAGTGCAGAA 240
QY 431 GGTGTTGGGCTGGAGGAATATCTGTCAGTGTGGTGGATGAGTCTCTCGGGCACCAC 490
Db 241 GGTGTTGGGCTGGAGGAATATCTGTCAGTGTGGTGGATGAGTCTCTCGGGCACCAC 300
QY 491 CAAAGTGCAGAGGCTCATCCGATAGCAGATTGGCAAGAACTTTCAGCAGAGAGCC 550
Db 301 CAAAGTGCAGAGGCTCATCCGATAGCAGATTGGCAAGAACTTTCAGCAGAGAGCC 360
QY 551 GCTGCGCAAGGAGGTGTCTCTGAGCAGGGGATCTGTTGCTGCTGCTGCTGCTGCTGCTG 610
Db 361 GCTGCGCAAGGAGGTGTCTCTGAGCAGGGGATCTGTTGCTGCTGCTGCTGCTGCTGCTG 420
QY 611 CATCTCTCCAGCGAGGTGGAGTGGTCCGGAACGAGGACCTGGTGGACCCGCTCCCTGGA 670
Db 421 CATCTCTCCAGCGAGGTGGAGTGGTCCGGAACGAGGACCTGGTGGACCCGCTCCCTGGA 480
QY 671 CCCCAATGTATACATCACCGGGAGCAGACCTGGTGGTGGTGGTGGTGGTGGTGGTGGTGG 730
Db 481 CCCCAATGTATACATCACCGGGAGCAGACCTGGTGGTGGTGGTGGTGGTGGTGGTGGTGG 540
QY 731 CACGGCCAACTACACCTGCTGGCCGCAAGAACTCTGTCGACCTGTCGCGCAGGCGCTCCGC 790
Db 541 CACGGCCAACTACACCTGCTGGCCGCAAGAACTCTGTCGACCTGTCGCGCAGGCGCTCCGC 600
QY 791 TGTGTCATCTGCTACGTGAAACGTTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGG 849
Db 601 TGTGTCATCTGCTACGTGAAACGTTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGG 659
QY 850 GCGCCAGCTGTGGGCGCGCTGGCAGAAACGAGGACCTGTCACCAACCCCGCGCGCTC 909
Db 660 GCGCCAGCTGTGGGCGCGCTGGCAGAAACGAGGACCTGTCACCAACCCCGCGCGCTC 719
QY 910 TCACCGGGGCGCTTTCTGTAGGGGAGAAATGTCATGACCGGACCGCTCTCTCTCTGCG 969
Db 720 TCACCGGGGCGCTTTCTGTAGGGGAGAAATGTCATGACCGGACCGCTCTCTCTCTGCG 774
QY 970 TTGTCCTGTGGAGCGCAGCTGGAGCCGCTGGAGCAGAGTGGTGGGCTGTGCTGCTGCTG 1027
```



```

Db      775 CTGTGCCAATGACGACGACTGTAGCAGCTGGAGCCAGTGGTGGGCTGTGGCTTGGC 834
QY      1028 CTGACCCACT--GGCGGACCGGTGAGTGTCTGACCCAGCACCACCGCAGCGAGGAG 1085
Db      835 TTGACCCACTTGGGCGGAGCCGAGTGTCTGAACCCAGCACCACCGCAGCGAGGAG 894
QY      1086 GAGTG---CCAGGGCACTGACCTGGACACCCCGCACTGTACCAGTGAACCTCTGTGTACAC 1142
Db      895 AGTGTGCGCCAGTCACTGGACCTGGACCCCGGA-TGGTCCAGTGAAGTCTGTGT-CCC 952
QY      1143 AGTGTCTTCTGGCCCTGAGGACGTCGGCCCTCTATGTGGGCTCATGCCG 1191
Db      953 ACGGGTTCTGGCCCTGAGGACTTGGCTCTCTATGTGGGCTCATGCCG 1001

RESULT 8
AY411749
LOCUS      2532 bp DNA linear GSS 12-DEC-2003
DEFINITION Mus musculus HCM4327 gene, VIRTUAL TRANSCRIPT, partial sequence,
AY411749
ACCESSION AY411749
VERSION
KEYWORDS
SOURCE
ORGANISM
Mus musculus (house mouse)
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
REFERENCE 1 (bases 1 to 2532)
AUTHORS Clark,A.G., Glanowski,S., Nielson,R., Thomas,P., Kejarawal,A.,
Todd,M.A., Tanenbaum,D.M., Civello,D.R., Lu,F., Murphy,B.,
Ferriera,S., Wang,G., Zheng,X.H., White,T.J., Sninsky,J.J.,
Adams,M.D. and Cargill,M.
TITLE Inferring nonneutral evolution from human-chimp-mouse orthologous
gene trios
JOURNAL Science 302 (5652), 1960-1963 (2003)
PUBMED 14671302
REFERENCE 2 (bases 1 to 2532)
AUTHORS Clark,A.G., Glanowski,S., Nielson,R., Thomas,P., Kejarawal,A.,
Todd,M.A., Tanenbaum,D.M., Civello,D.R., Lu,F., Murphy,B.,
Ferriera,S., Wang,G., Zheng,X.H., White,T.J., Sninsky,J.J.,
Adams,M.D. and Cargill,M.
TITLE Direct Submission
JOURNAL Submitted (16-NOV-2003) Celera Genomics, 45 West Gude Drive,
Rockville, MD 20850, USA
COMMENT This sequence was made by sequencing genomic exons and ordering
them based on alignment.
FEATURES
source Location/Qualifiers
1..2532
/organism="Mus musculus"
/mol_type="genomic DNA"
/db_xref="taxon:10090"
<1..>2532
/locus_tag="HCM4327"

ORIGIN
Query Match 25.2%; Score 726.4; DB 29; Length 2532;
Best Local Similarity 59.1%; Pred. No. 1.2e-111;
Matches 1492; Conservative 0; Mismatches 896; Indels 138; Gaps 9;

QY      393 GAGTCCGCATTATGTCTCAGGACGACGAGTCGAGAGGTGTTGGGCTGGAGGAATAC 452
Db      13 GAGTGCAGATCGAGGTGTACGGCAGCAGTGGAGGAATCTTTGGGCTCGAGGACTAC 72
QY      453 TGGTGCAGTGCCTGGATGAGTCTCTCGGACACCGACCGAGTTCAGAGGCTTACATC 512
Db      73 TGGTGCAGTGCCTGGCTGAGTCTTCGGAACTACCAAGAGTCCCGGACCTACATC 132
QY      513 CGCATGCCAATTGCGCAAGAACTTCGACGAGGCGCTGGCCAAAGGAGTGTCCCTG 572
Db      133 CGCATTCCTTACTTGGCGAAGAACTTTGACGAGAGCTCTGGCCAAAGGAGTACCTTG 192
QY      573 GAGCAGGCGATCGTGTGCCCTGCGCTCCACCGGAGGCGATCCCTCCAGCGGAGTGGAG 632

```

```

Db      193 GATCATGAGGTCTTCTGTGAGTCCGCCCCACCGAGGGAGTGCCTGTGGCTGAGGTGGAA 252
QY      633 TGCTCCGGAACGAGGACCTTGTGTGACCCGTCCCTGTGAGCCCAATGTATACATCATCCGCG 692
Db      253 TGCTCAAGAAATGAAGATGTCTATTGACCCCGCTCAGACACATACTTCTGTCTCACCATT 312
QY      693 GAGCAGAGCTGGTGGTGGCAGAGGCGCCCTTGTGTGACGAGGCGCAACTACACTCGGTG 752
Db      313 GACCAACACTCATCATCCGCGAGGCGCGCTCTCAGACACGCGCAACTACACTGTGTG 372
QY      753 GCCAAGAACATCGTGGCAGCTCGCGCAGCGCTCGGCTGTGTGATCTGTCTGTCTGTGTAAC 812
Db      373 GCCAAGATATCGTGGCAGCGCGGAGCACACCGCCACAGTCACTGTCTGTGTGAAT 432
QY      813 GGTGGTGTGTGACCTGTGGACCGAGTGGTCCGTCTGTGAGCGCAGCTGTGGGCGCGGCTGG 872
Db      433 CGAGGCTGTGTGAGTGGGACAGTGGTCACTGTGTGCGCTGTGCGCGAGGCTGG 492
QY      873 CAGAAACGAGCGCGAGCTGTGACCAACCGGCGCTCTCAACGAGGCGGCTTTCTGTGTGAG 932
Db      493 CAGAAAGCTGTCTGAGCCTGCACTCAATCCAGCCCACTCAATGAGAGGCGCTTCTGTGTGAG 552
QY      933 GGGCAGAAATGTCCATGACCGCAGCCCTCTCTCTGTGTGTGTCTGTGTGACGCGAGCTGG 992
Db      553 GGACAGGCTTCCAGAA---GACAGCTTGCAACCGTGTGCGCAGTGGATGGAGCGTGG 609
QY      993 AGCCGCTGGAGCAAGTGTGGGCTGTGGGTGGACTGCACCCACTTGGCGGAGCGGTGTGAG 1052
Db      610 ACCGAGTGGAGCAAGTGTGTCTGTCTGTGACGACAGAGTGTGCGCACTTGGCGCAGCGCGAG 669
QY      1053 TGCTCTGACCCAGCAGCAGCGCAACCGAGGAGGAGTGTGAGGCGCACTGTGACGACACC 1112
Db      670 TGATGGCAGCGCCACCCAGAACAGGCGCGTGTGATGACGCGGAGCGCTACTTCACTCC 729
QY      1113 CGCAACTGTACAGTGAACCTGTGTGTACACAGTGTCTTGTGCGCTGTGAGGAC--- 1163
Db      730 AAGAACTGCACTGATGGCTGTGGTGTGTGATCAGAACTCTTAAACGACCTTAAAGC 789
QY      1164 -----GTGGCCCTCTATGTGGGCTCATCGCGTGGCGGCTC 1199
Db      790 CACCCCTGGAGACATCGGAGATGTGCACTGTACGCGAGGCTTGTGTGCGCTCTTT 849
QY      1200 TGCTGTGCTCTGTCTGTCTGTCTCT---CATCTCTGTTTATGCGGAGAGAGAGGG 1256
Db      850 GTGGTGGTGGGCTGTCTCATGGCGTGGAGTGTATCGTATACCGGAGAACTTGC CGGAC 909
QY      1257 CTGACTCAGATGTGGTGTGACTCGT---CCATTTCTCACTCAGGCTTCCAGCGCTCAGC 1313
Db      910 TTCGACACGAGACATCACCGACTCTCTGTGGGCTCTACTGTGTGCTTCCACCTGTCAAC 969
QY      1314 ATCAAGCCCGCAAGCAGACACCCCATATGTGT-----CACCATCCAGCGGACCTC 1367
Db      970 TTCAAGACTGCAAGGCGCAACCAACCCGCGAGTCTCTGCACCGCTCGCGCCCTCCAGACCTA 1029
QY      1368 AGCACCACCACTTACCGAGGAGTCTGTGTCCCGGAGGATGGGCGGAGGAGGAG 1427
Db      1030 ACGGCCAGTGTGTGCACTTACCGCGGCTGTGTATGCCCTGTGAGGACTTCGCGGACAAG 1089
QY      1428 TTCCAGCTCACCAATGGGCACCTGTCT----- 1453
Db      1090 ATCCCATGACTAATTCGCCCTTGTGGATGCCCTGCCAGCTCAAGATCAAGGTCTAT 1149
QY      1454 -----CAGCCCTCTGGGTGGC 1469
Db      1150 AACTCCAGCACCATTCTGGTCTGGCTGGCTGGTGTGATGGAGCGGACTGTCTGGGTCTC 1209
QY      1470 GCGCGGCACACACTGACACAGCTCTCCACCTCTGAGGCGGAGGAGTGTGTCTCCCGC 1529
Db      1210 CTCGCGCGGCGACGTACCCAGGCGATTTCTCCGGGACACCCATTTCTGTGCACTGCGC 1269
QY      1530 CTCTCCACCAGAACT-----ACTTCGCTCTCCCGGAGGACCAACACATGACC 1583
Db      1270 AGTGCCAGCCTTGGTTCCTCCAGCACTCTCTGGGCTTACCTCGGAGCCCGGAGCAGTGTCT 1329

```

```
QY 1584 TATGGACCTTTCAATCTCTCGGGGCGCGGTGATGATCCCTAATACAGGTATCAGCCTC 1643
Db 1330 AGCGGACCTTTGGTTGGCTGGAGGAAGGTGAGCTCCCGGGCACAGGGGTGAGCCTG 1389
QY 1644 CTATATCCCCCAGATGCCATACCCCGAGGGAAGATCTATAGATCTACCTCAGCGTGCAC 1703
Db 1390 TTGTTACCAATGGAGCCATTTCCCGAGGCGAGTTCTATGACCTGTATCTACATATCAAC 1449
QY 1704 AAGCGGAAGACGTGAGGTGGCCCTAGCTGTGCTGCTCAGACCCCTGCTGAGTCCCATCGTT 1763
Db 1450 AAGGCGGAAGACACCTCCCATTTTCAGAGGTTCCAGACAGTATTGAGCCCTCGGTG 1509
QY 1764 AGCTGTGACCCCTGGCGTCTGCTACCCCGGCAGTCACTCTGCTGATGACCACTGT 1823
Db 1510 ACCTGTGGGCGCACAGGCTACTCTCTGTGCGCGCTGTCTCTCACCCTGCGCCACTGT 1569
QY 1824 GGGAGCCAGCCCTGACAGCTGGAGCTGCGCTCAAAAAGCAGTCTGTGGAGGGCAGC 1883
Db 1570 GCTGAAGTATCGCTGGAGACTGGATCTTTAGCTCAAGACCCAGGCCCATCAGGGCCAC 1629
QY 1884 TGGAGCAGGATGTCTGCACCTGGCGGAGGAGGCGCCCTCCACCTCTACTACTGCCAG 1943
Db 1630 TGGG---AGGAGTGTGACCTTGGATGAGGAGACCCCTCAACACACCTGTCTACTGCCAG 1686
QY 1944 CTGGAGGCGAGTGTCTGCTTTCACCGAGAGCTGGGCGGCTTGGCCCTGTGGGA 2003
Db 1687 CTGGAGGCTAAGTCTCGCCACATCTCTGTGGACCACTGGGTACCTAGTATTCATGGGC 1746
QY 2004 GAGGCGCTCAGGCTGGCTGCCCGCAGCGCTCAAGCTGCTTCTGTTGCGCGGTGGCC 2063
Db 1747 GAGTCTTACTCTCGTCTGCACTGAAGCGGCTCCAGCTGGCCATCTTGGCCCGAGCCTC 1806
QY 2064 TGCACCTCCCTCGAGTACACATCCGGGTCTACTGCTGCATGACACCCACGATGCATC 2123
Db 1807 TGCACTCCCTGAGTATAGCTCAGGCTCTACTGCTGGAGGACACACCTGTAGCAGT 1866
QY 2124 AAGGAGGTGGTGAAGCTGAGAGAGCTGGGGGAGACGCTGATCCAGAGGACACGGTTC 2183
Db 1867 AAGGAGGTCTGAGCTGGAGAGGACTCTGGGTGGCTACTTGGTGGAGGAGGCCAAGCCT 1926
QY 2184 CTGCACTTCAAGGACAGTTTACACACCTCGCGCTATCCATCCAGCATGTGCCAGCTCC 2243
Db 1927 TTGCTCTTTAAGACAGTTTACACACCTTACAGCTTCCCTCCATGACATCCCATGCC 1986
QY 2244 CTGTGGAAGAGTAAGCTCCCTGTGCTACCTACAGGAGATCCCTTTTATACATCTGGAAT 2303
Db 1987 CACTGGAGGAGCAAACTACTGGCCAAAGTACAGGAGATTCCCTTCTPACCACTGTGGAAT 2046
QY 2304 GGCAGCAGGGTACTTGCTACTGCTACCTTCACTGGAGGTGTGAGCCCGACACTAGT 2363
Db 2047 GGCAGCAGAGAGCCCTGCACTGCACTTTTACCTGGAGGAGGATAGCTTGGCCTCCAG 2106
QY 2364 GACCTGGCTGCAAGCTGTGGGTGTGGCAGGTGGAGGCGACGGGAGAGCTTCAGCATC 2423
Db 2107 GAGTTCACCTGTAAAGTCTGCTGCGGAGGCTCAAGGGAAGGCCAGATTTTCCAGCTG 2166
QY 2424 AACTTCAACATCACCC---AAGGACACAAAGTTTGTGCTGAGCTGTGCTCTGAGAGTGAA 2480
Db 2167 CACACACAGTTTGGCCGAGACGCTGTGCTGCTCCCTGGATGCTCTCTGCTGCGCCGGCC 2226
QY 2481 GCGGGGTCCAGCCTGTGGTGGGCGCCAGTGCCTTCAAGATCCCTTCTCATTTCCGCGAG 2540
Db 2227 AATGCCATCACCCAGCTGGAGCCCTATGCTTCAAGATACCCCTGTCTCATTCGCCGCAA 2286
QY 2541 AAGATATTTCCAGCTTGGACCCACCTGTAGGCGGGGTGCCGACTGGCGACTCTGGCC 2600
Db 2287 AAGATCTGAGAGGCTTGGAGCGCCCCCACTCCGGGGCAACGACTGGAGGCTGTGGCG 2346
QY 2601 CAGAAATCCCACTGGACAGCCATCTCAGCTTCTTGGCTTCCAGGCCAGCCCAAGCC 2660
Db 2347 CAGAAGCTGTGTCATGGACCGGTACTTAACTACTTGGCCACCAAGAGTACTGCCACAGGT 2406
```

```
QY 2661 ATGATCTCAACTGTGGAGGCGCGGCACCTTCCCAACGGCAACCTCAGCCAGCTGGCT 2720
Db 2407 GTCACTTAGACCTCTGGAGCTCGCAACAGAGTACGGGACCTCAACAGCCTGGCC 2466
QY 2721 GCACAGTGGCTGACTGGCGGACGACGCTGCTCTTTCACAGTCTCGAGGCTGAG 2780
Db 2467 AGTCCTTGGAGGAGATGGCAAGAGTGAGATGCTGTGCTAGCCATGGCCACAGATGGCAT 2526
QY 2781 TGCTGA 2786
Db 2527 TGCTGA 2532

RESULT 9
BX364574 309 bp mRNA linear EST 05-MAY-2003
LOCUS BX364574 Homo sapiens NEUROBLASTOMA COT 25-NORMALIZED Homo sapiens
DEFINITION CDNA clone CS0DC023YB10 5-PRIME, mRNA sequence.
ACCESSION BX364574
VERSION BX364574.1 GI:30368812
KEYWORDS EST.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
REFERENCE 1 (bases 1 to 909)
AUTHORS Li, W.B., Gruber, C., Jessee, J. and Polayes, D.
TITLE Full-length cDNA libraries and normalization
JOURNAL Unpublished (2001)
COMMENT Contact: Genoscope
Genoscope - Centre National de Sequencage
BP 191 91006 EVRY cedex - France
Email: seqre@genoscope.cns.fr, Web: www.genoscope.cns.fr
Library was constructed by Life Technologies, a division of
Invitrogen. This sequence belongs to sequence cluster 3529.f For
more information about this cluster, see
http://www.genoscope.cns.fr/cgi-bin/cluster.cgi?seq=CS1AC006ZF06P1&cluster=3529.f. Contact:
Peng Liang Email: fliang@lifetech.com URL:
http://fulllength.invitrogen.com/ Invitrogen Corporation 1600
Paradise Avenue Genoscope sequence ID: CS1AC006ZF06QP1.

FEATURES
Location/Qualifiers
1..909
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="CS0DC023YB10"
/tissue_type="NEUROBLASTOMA COT 25-NORMALIZED"
/clone_lib="Homo sapiens NEUROBLASTOMA COT 25-NORMALIZED"
/note="1st strand cDNA was primed with a NotI-oligo(dT)
primer. Five prime end enriched, double-strand cDNA was
digested with Not I and cloned into the Not I and EcoR V
sites of the pCMVSPORT 6 vector. Library was normalized."

ORIGIN
Query Match 25.2%; Score 726.2; DB 13; Length 909;
Best Local Similarity 98.4%; Pred. No. 7,9e-112;
Matches 739; Conservative 4; Mismatches 7; Indels 1; Gaps 1;

QY 1 AGTGGGGTCCGGGCTGAGCGCTAAAGCGCCCTCCCGCGGGGGGGGGGGGGGGGGGGGG 60
Db 160 AGTGGGGTCCGGGCTGAGCGCTAAAGCGCCCTCCCGCGGGGGGGGGGGGGGGGGGG 219
QY 61 GCGCGCGCGCTCCCGCGGGGCGCATGCGCTCGCGCGCGCGCGCTGTGCGCAGCGCTCC 120
Db 220 GCGCGCGCGCTCCCGCGGGGCGCATGCGCTCCCGCGCGCGCGCTGTGCGCAGCGCTCC 279
QY 121 TGGGCAATAGTCTTCGCGGCTTGGCTCCGGGCTGCGGCTGCGCGCGCGCTGTGCGCAGCGCTGG 180
Db 280 TGGGCAATAGTCTTCGCGGCTTGGCTCCCGGCTCGGGTGGCGCGCGCGCTGTGCGCAGCGCTGG 339
QY 181 CCAACCAAGTGTGTGTGCGCAACCGCGGACTGTCTTCCCACTTCTCTGGTGGAGCCGAGG 240
```


RESULT 11

BC033727

LOCUS

DEFINITION

ACCESSION

VERSION

KEYWORDS

SOURCE

ORGANISM

REFERENCE

AUTHORS

BC033727 1532 bp mRNA linear HTC 19-NOV-2003
Homo sapiens netrin receptor Unc5h1, mRNA (cDNA clone
IMAGE15166762), containing frame-shift errors.

BC033727.1 GI:21707230

HTC.

Homo sapiens

Homo sapiens (human)

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.

1 (bases 1 to 1532)

Klausner, R.D., Collins, P.S., Wagner, L., Shenmen, C.M., Schuler, G.D.,

Altschul, S.F., Zeeberg, B., Buetow, K.H., Schaefer, C.F., Bhat, N.K.,

Hopkins, R.F., Jordan, H., Moore, T., Max, S.I., Wang, J., Hsieh, P.,

Diatchenko, L., Marusina, K., Farmer, A.A., Rubin, G.M., Hong, L.,

Stapleton, M., Soares, M.B., Bonaldo, M.F., Casavant, T.L.,

Scheetz, T.E., Brownstein, M.J., Usdin, T.B., Toshiyuki, S.,

Carninci, P., Prange, C., Raha, S., Loquellano, N.A., Peters, G.J.,

Abrahamson, R.D., Mullany, S.J., Bosak, S.A., McEwan, P.J.,

McKernan, K.J., Malek, J.A., Gunaratne, P.H., Richards, S.,

Worley, K.C., Hale, S., Garcia, A.M., Gay, L.J., Hulyk, S.W.,

Villalon, D.K., Muzny, D.M., Sodergren, E.J., Lu, X., Gibbs, R.A.,

Fahy, J., Helton, E., Kettman, M., Madan, A., Rodrigues, S.,

Sanchez, A., Whiting, M., Madan, A., Young, A.C., Shevchenko, Y.,

Bouffard, G.G., Blakesley, R.W., Touchman, J.W., Green, E.D.,

Dickson, M.C., Rodriguez, A.C., Grimwood, J., Schmutz, J., Myers, R.M.,

Butterfield, Y.S., Krzywinski, M.I., Skalska, U., Smallos, D.E.,

Schmerch, A., Schein, J.E., Jones, S.J. and Marra, M.A.

Generation and initial analysis of more than 15,000 full-length

human and mouse cDNA sequences

Proc. Natl. Acad. Sci. U.S.A. 99 (26), 16899-16903 (2002)

2238257

2247732

2 (bases 1 to 1532)

Strausberg, R.

Direct Submission

Submitted (02-JUL-2002) National Institutes of Health, Mammalian

Gene Collection (MGC), Cancer Genomics Office, National Cancer

Institute, 31 Center Drive, Room 11A03, Bethesda, MD 20892-2590,

USA

NIH-MGC Project URL: <http://mgc.nci.nih.gov>

Contact: MGC help desk

Email: cgabs-r@mail.nih.gov

Tissue Procurement: Life Technologies, Inc.

cDNA Library Preparation: Life Technologies, Inc.

DNA Sequencing by: The I.M.A.G.E. Consortium (LLNL)

Sequencing Center (NISC)

Gaithersburg, Maryland;

Web site: <http://www.nisc.nih.gov/>

Contact: nisc_mgc@nih.gov

Akhter, N., Ayele, K., Beckstrom-Sternberg, S.M., Benjamin, B.,

Blakesley, R.W., Bouffard, G.G., Green, K., Brinkley, C., Brooks, S.,

Dietrich, N.L., Granite, S., Guan, X., Gupta, J., Haghighi, P.,

Hansen, N., Ho, S.-L., Karlins, E., Kwong, P., Laric, P., Legaspi, R.,

Maduro, O.L., Masiello, C., Maskeri, B., Mastrian, S.D., McCloskey, J.C.,

McDowell, J., Pearson, R., Stantripop, S., Thomas, P.J., Touchman, J.W.,

Thurgeon, C., Vogt, J.L., Walker, M.A., Wetherby, K.D., Wiggins, L.,

Young, A., Zhang, L.-H. and Green, E.D.

Clone distribution: MGC clone distribution information can be found

through the I.M.A.G.E. Consortium/LLNL at: <http://image.llnl.gov>

Series: IRAC Plate: 68 Row: i Column: 2

This clone was selected for full length sequencing because it

passed the following selection criteria: GenomeScan gene prediction

This clone has the following problem: frame shifted.

Location/Qualifiers

1..1532

/organism="Homo sapiens"

source

FEATURES

ORIGIN

Query Match

Best Local Similarity

Matches 910; Conservative

24.7%; Score 711.2; DB 11; Length 1532;

83.3%; Pred. No. 3.3e-109; Indels 174; Gaps 2;

Mismatches 8; Indels 174; Gaps 2;

/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="IMAGE:5166762"
/tissue_type="Brain, adult medulla"
/clone_lib="NIH_MGC_l19"
/lab_host="DH10B"
/note="Vector: pCMV-SPORT6"

ORIGIN

378 GGTGAGCGGACCATCGAGTCCGATTATATGTTCTCAAGCAGCAGAGTTCGAGAAGTGTTC 437
612 GGGTGGCCACCATGAGAGTCCGATTATATGTTCTCAAGCAGCAGAGTTCGAGAAGTGTTC 671
438 GGGCTGGAGGAATACTGTGTCAGTGGCGATGGAGTCTCTGGGACACCAAGAGT 497
672 GGGCTGGAGGAATACTGTGTCAGTGGCGATGGAGTCTCTGGGACACCAAGAGT 731
498 CAGAGGCTTACATCCGATAGCAGATTGGCAGAGACTTCGAGCAGGAGCGCTGGCC 557
732 CAGAGGCTTACATCCGATAGCAGATTGGCAGAGACTTCGAGCAGGAGCGCTGGCC 791
558 AAGGAGGTGTCCCTCGAGCAGGAGCATCGTGTGTCCTCCACCGAGGAGTCCCT 617
792 AAGGAGGTGTCCCTCGAGCAGGAGCATCGTGTGTCCTCCACCGAGGAGTCCCT 851
618 CCAGCGGAGTGGAGTGGCTCCGGAACGAGAGACTGTGTGGACCCGCTCCGTGACCCCAAT 677
852 CCAGCGGAGTGGAGTGGCTCCGGAACGAGAGACTGTGTGGACCCGCTCCGTGACCCCAAT 911
678 GTATACATCAGCGGAGCAGACGCTGTGTGGAGCAGCGCCCTTGTGACACGGCC 737
912 GTATACATCAGCGGAGCAGACGCTGTGTGGAGCAGCGCCCTTGTGACACGGCC 971
738 AACTACACCTCGTGGCCAAAGAACATCGTGGCAGCAGTTCGCGCAGCGCCCTCCGTGTGTC 797
972 AACTACACCTCGTGGCCAAAGAACATCGTGGCAGCAGTTCGCGCAGCGCCCTCCGTGTGTC 1031
798 ATCGTCTACGTGAACGGTGGTGGTGGAGCTGAGCAGTGGAGCGAGTGTGCTGTGACGCGCAGC 857
1032 ATCGTCTAC----- 1040
858 TGTGGCGCGGCTGGCGAGAAACGAGCAGGAGTGCACCAACCGCGCGCTCTCAACGGG 917
1041 ----- 1040
918 GGGCGCTTCTGTGAGGGGAGAAATGTCATGACCGCAGCCGTCTCTCTGTGTTGTCTCT 977
1041 ----- 1040
978 GTGAGCGGAGCTGGAGCGCGCTGGAGCAAGTGTGCGCCTGTGGGCTGGACTGCACCCAC 1037
1041 GTGAGCGGAGCTGGAGCGCGCTGGAGCAAGTGTGCGCCTGTGGGCTGGACTGCACCCAC 1100
1038 TGGCGGAGCGCTGAGTGTCTGACCCAGCAGCAGCCGCAACCGAGGAGGAGTGCACAGGC 1097
1101 TGGCGGAGCGCTGAGTGTCTGACCCAGCAGCAGCCGCAACCGAGGAGGAGTGCACAGGC 1160
1098 ACTGACCTGGACACCGCGCAACTGTACAGTGAACCTCTGTGTACACAGTGTCTTGGCCCT 1157
1161 ACTGACCTGGACACCGCGCAACTGTACAGTGAACCTCTGTGTACACAGTGTCTTGGCCCT 1220
1158 GAGGAGCTGGCCCTCTATGTGGGCTCATCGCGTGGCGCTTGCCTGCTGCTGCTGCTG 1217
1221 GAGGAGCTGGCCCTCTATGTGGGCTCATCGCGTGGCGCTTGCCTGCTGCTGCTGCTG 1280
1218 CTGTGCTCATCTCTGTTTATTTCGGAAGAGAGGAGGCTGGAGTGCAGATGTGGCTGAC 1277
1281 CTGTGCTCATCTCTGTTTATTTCGGAAGAGAGGAGGCTGGAGTGCAGATGTGGCTGAC 1340

```

QY 1278 TCGTCCATTCACCTCAGGCTCCAGCCCGTCAGCATCAAGCCCAAGCAGACAAC 1337
Db 1341 TCGTCCATTCACCTCAGGCTCCAGCCCGTCAGCATCAAGCCCAAGCAGACAAC 1400
QY 1338 CCCATCTGTCAGCATCAGCCGACCTCAG---CACCAACCAACCTACCAAGGCGAGT 1394
Db 1401 CCCATCTGTCAGCATCAGCCGACCTCAGCACCACCAACCAACCTACCAAGGCGAGT 1460
QY 1395 CTCTGTCCCGGCGAGGATGGCCCGAGCCCAAGCTTCAGCTCAACATGGGCACTGTCTC 1454
Db 1461 CTCTGTCCCGGCGAGGATGGCCCGAGCCCAAGCTTCAGCTCAACATGGGCACTGTCTC 1520
QY 1455 AGCCCTCTGGGT 1466
Db 1521 AGCCCTCTGGGT 1532

RESULT 12
LOCUS BI818609
DEFINITION 60303362F1 NIH_MGC_115 Homo sapiens cDNA clone IMAGE:5174559 5',
mRNA sequence.
ACCESSION BI818609
VERSION BI818609.1 GI:15929902
KEYWORDS EST.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
NIH-MGC http://mgi.nci.nih.gov/.
National Institutes of Health, Mammalian Gene Collection (MGC)
Unpublished (1999)
Contact: Robert Strausberg, Ph.D.
Email: cgapbs-r@mail.nih.gov
Tissue Procurement: Life Technologies, Inc.
cDNA Library Preparation: Life Technologies, Inc.
cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
DNA Sequencing by: Incyte Genomics, Inc.
Clone distribution: MGC clone distribution information can be
found through the I.M.A.G.E. Consortium/LLNL at:
http://image.llnl.gov
Plate: LLAM11434 row: 1 column: 16
High quality sequence stop: 744.
FEATURES
    source
        1..818
            /organism="Homo sapiens"
            /mol_type="mRNA"
            /db_xref="taxon:9606"
            /clone="IMAGE:5174559"
            /lab_host="DH10B"
            /clone_lib="NIH_MGC_115"
            /note="Organ: pooled brain, lung, testis; Vector:
            pCMV-Sport6; Site 1: NotI; Site 2: EcoRV (destroyed); RNA
            source anonymous pool of 6 male brains, age range 23-27; 1
            male lung, age 27; and 1 male testis, age 69. Library is
            oligo-dT primed and directionally cloned (EcoRV site is
            destroyed upon cloning). Average insert size 1.8 kb,
            insert size range 1-3 kb. Library is normalized and
            enriched for full-length clones and was constructed by C.
            Gruber (Invitrogen). Research Genetics tracking code
            021. Note: this is a NIH_MGC Library."

Query Match 24.6%; Score 710; DB 12; Length 818;
Best Local Similarity 97.1%; Pred. No. 4e-109;
Matches 744; Conservative 0; Mismatches 20; Indels 2; Gaps 2;

QY 84 GCCATGGCCGTCGGCGCGGCTGTGGCCAGCGCTCTGGGCATAGCTCCGCGCTTGG 143
Db 1 GCCATGGCCGTCGGCGCGGCTGTGGCCAGCGCTCTGTGGGCATAGCTCCGCGCTTGG 60
QY 144 CTCGCGGGCTCGGTGTCGCCAGAGGTGCCACCGTGGCCACCAACCAAGTGCCTGTGTCAC 203

```

```

Db 61 CTCGCGGGCTCGGGTCCCGCAGCAGAGTGCCACCGTGGCCCAACCCAGTGCCTGTGCGCAAC 120
QY 204 CCGGACCTGCTTCCCTCCACTTCTCTGTTGGAGCCCGCAGGATGTGTACATCGTCAAGACAAG 263
Db 121 CCGGACCTGCTTCCCTCCACTTCTCTGTTGGAGCCCGCAGGATGTGTACATCGTCAAGACAAG 180
QY 264 CCAAGTCTGCTTGTGTGCAAGGCCGTGCCCGCCACACAGATCTTCTTCAAGTGCACAGGG 323
Db 181 CCAAGTCTGCTTGTGTGCAAGGCCGTGCCCGCCACACAGATCTTCTTCAAGTGCACAGGG 240
QY 324 GAGTGGTGGCCAGCTGGACACACAGTGTGAGCGGACAGACAGACGCGAGCAGTGTGTAG 383
Db 241 GAGTGGTGGCCAGCTGGACACACAGTGTGAGCGGACAGACAGACGCGAGCAGTGTGTAG 300
QY 384 CCGACCATGGAGTCCGCAATTAATGCTCAAGCGACAGCTCGAGAAGGTGTCGGGGCTG 443
Db 301 CCCACCATGGAGTCCGCAATTAATGCTCAAGCGACAGTTCAGAAGGTGTCGGGGCTG 360
QY 444 GAGGAATCTGTGTCAGTGCCTGGCATGGAGTCTCTCGGGCACCACCAAGAGTCAAGAAG 503
Db 361 GAGGAATCTGTGTCAGTGCCTGGCATGGAGTCTCTCGGGCACCACCAAGAGTCAAGAAG 420
QY 504 GCTATCATCCGCATAGCCAGATTGGCAAGAACTTCGAGCAGGAGCCGCTGCCCAAGAG 563
Db 421 GCTATCATCCGCATAGCCCTTAATTTGGCAAGAACTTCGAGCAGGAGCCGCTGCCCAAGAG 480
QY 564 GTGTCCTGGAGCAGCGCATCTGTGTCCTCCCTGCCCTCCACCGAGGGCATCCCTCCAGCC 623
Db 481 GTGTCCTGGAGCAGCGCATCTGTGTCCTCCCTGCCCTCCACCGAGGGCATCCCTCCAGCC 540
QY 624 GAGTGGAGTGGCTCCGGAACGAGGACCTGTGTGGACCCGCTCCCTGGACCCCAATGTATAC 683
Db 541 GAGTGGAGTGGCTCCGGAACGAGGACCTGTGTGGACCCGCTCCCTGGACCCCAATGTATAC 600
QY 684 ATCAGCGGGAGCAGACGCTGTGTGTGACAGGCCCGCTTGTGTGACAGGCCCAACTAC 743
Db 601 ATCAGCGGGAGCAGACGCTGTGTGTGACAGGCCCGCTTGTGTGACAGGCCCAACTAC 660
QY 744 ACCTGCGTGGCCAGAACATCTGTGACAGTC-GGCGAGCGCTCCGCTGCTGTGTATCGT 802
Db 661 ACCTGCGTGGCCAGAACATCTGTGACAGTC-GGCGAGCGCTCCGCTGCTGTGTATCGT 720
QY 803 CTAGTGAACGCTGGGTGGTTCGACG-TGGACCGAGTGGTTCGCTGTG 847
Db 721 CTAGTGAACGCTGGGTGGTTCGACG-TGGACCGAGTGGTTCGCTGTG 766

RESULT 13
EX348193/c
LOCUS BX348193
DEFINITION BX348193 Homo sapiens NEUROBLASTOMA COT 10-NORMALIZED Homo sapiens
cDNA clone CS0DB008YE02 5-PRIME, mRNA sequence.
ACCESSION BX348193
VERSION BX348193.1 GI:30367259
KEYWORDS EST.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
Li,W.B., Gruber,C., Jesse,J. and Polayes,D.
1.(bases 1 to 843)
Full-length cDNA libraries and normalization
Unpublished (2001)
Contact: Genoscope
Genoscope - Centre National de Sequencage
BP 191 91006 EVRY cedex - France
Email: seqref@genoscope.cns.fr, Web : www.genoscope.cns.fr
Library was constructed by Life Technologies, a division of
Invitrogen. This sequence belongs to sequence cluster 3239.r For
more information about this cluster, see
http://www.genoscope.cns.fr/
cgi-bin/cluster.cgi?seq=CS0BAF004ZD01_AF00293_1kcluster=3239.r.

```

Contact : Feng Liang Email : fliang@lifetech.com URL :
 http://fulllength.invitrogen.com/ Invitrogen Corporation 1600
 Faraday Avenue Genoscope sequence ID : CS0BAF004ZB01_AF00293_1.
 Location/Qualifiers

FEATURES

source

1. .843
 /organism="Homo sapiens"
 /mol_type="mRNA"
 /db_xref="taxon:9606"
 /clone="CS0DB008YB02"
 /tissue_type="NEUROBLASTOMA COT 10-NORMALIZED"
 /clone_lib="Homo sapiens NEUROBLASTOMA COT 10-NORMALIZED"
 /note="1st strand cDNA was primed with a NotI-oligo (dT)
 primer. Five prime end enriched, double-strand cDNA was
 digested with Not I and cloned into the Not I and EcoR V
 sites of the pCMVSPORT 6 vector. Library was normalized."

ORIGIN

Query Match 24.6%; Score 709.8; DB 13; Length 843;
 Best Local Similarity 95.0%; Pred. No. 4.3e-109;
 Matches 776; Conservative 0; Mismatches 34; Indels 7; Gaps 4;

1694 CACGCTGCACAGCCGGAAGAGCTGAGG--TTGCCCTTAGTGTGCTGTCAGACCCCTGTG 1751
 |||||
 824 CACGCGCCAAAGCGGAAGAGTGTAGGTTGCCCAACCCGGGTGTGACACCCCTGTG 765
 |||||
 1752 AGTCCATCTTAGTGTGAGCCCTCTGCGTCTGCTCACCGGCCAGTCTATCTGGCT 1811
 |||||
 764 AGTCCCATGTAGTGTGA--CCCTGGCGTCTGCTCACCGGCCAGTCTATCTGGGT 706
 |||||
 1812 ATGGACCACTGTGGGAGCCAGCCCTGACAGCTGGAGCTGCGCTC--AAAAAGCAGTC 1870
 |||||
 705 ATGGACCACTGTGGGAGCCAGCCCTGACAGCTGAGGCTGGCTCTCAAAAACAGTC 646
 |||||
 1871 GTGCGAGGCACTGGGAGCAGATGTGTGACCTTGGGCGAGAGCGCCCTCCACCT 1930
 |||||
 645 GTGCGAGGACAGCTGG--AGTATGCTGCACTTGGCGGAGNAGCGCCCTCCACCT 589
 |||||
 1931 CTACTACTCCAGCTGGAGCCAGTGCCTGCTACGCTCTTCCAGCAGCAGTGGCGGCTT 1990
 |||||
 588 CTACTACTCCAGCTGGAGCCAGTGCCTGCTACGCTCTTCCAGCAGCAGTGGCGGCTT 529
 |||||
 1991 TCCCTGTGGGAGAGCCCTCAGCGTGGTGGCGCAAGCGCTCAAGCTGCTCTGTT 2050
 |||||
 528 TGCCCTGTGGGAGAGCCCTCAGCGTGGTGGCGCAAGCGCTCAAGCTGCTCTGTT 469
 |||||
 2051 TGGCGCGGTGGCTGCACCTCCCTCGAGTACAACATCCGGGTCTATGCTGTGATGACAC 2110
 |||||
 468 TGGCGCGGTGGCTGCACCTCCCTCGAGTACAACATCCGGGTCTATGCTGTGATGACAC 409
 |||||
 2111 CCAGATGCACTCAAGAGGTGTGCGAGCTGGAGAGCAGCTGGGGGACAGCTGATCCA 2170
 |||||
 408 CCAGATGCACTCAAGAGGTGTGCGAGCTGGAGAGCAGCTGGGGGACAGCTGATCCA 349
 |||||
 2171 GGAGCAGCGGTCTCTGCACTTCAAGGACAGTTATCCACAACTCGCGCTATCCATCCACGA 2230
 |||||
 348 GGAGCAGCGGTCTCTGCACTTCAAGGACAGTTATCCACAACTCGCGCTATCCATCCACGA 289
 |||||
 2231 TGTGCCAGCTCCCTGTGGAAGAGTAAAGTCTCTTGTGAGTACAGAGATCCCTTTTA 2290
 |||||
 288 TGTGCCAGCTCCCTGTGGAAGAGTAAAGTCTCTTGTGAGTACAGAGATCCCTTTTA 229
 |||||
 2291 TCACATCTGGAATGGACGAGCGGTACTTGGCACTGCACTTCCACCTGGAGCGTGTGAC 2350
 |||||
 228 TCACATCTGGAATGGACGAGCGGTACTTGGCACTGCACTTCCACCTGGAGCGTGTGAC 169
 |||||
 2351 CCCAGCAGCTAGTGAACCTGGCCCTGCAAGCTGTGGGTGTGAGGTGGAGGCGCAGCGGCA 2410
 |||||
 168 CCCAGCAGCTAGTGAACCTGGCCCTGCAAGCTGTGGGTGTGAGGTGGAGGCGCAGCGGCA 109
 |||||
 2411 GAGCTTCAGCATCACTTCAACATCACCAGGACAGAGTTTGTCTGAGCTGTGGCTCT 2470
 |||||
 108 GAGCTTCAGCATCACTTCAACATCACCAGGACAGAGTTTGTCTGAGCTGTGGCTCT 49

QY 2471 GGAGAGTGAAGCGGGGTCCAGCCCTGTGTTGGGCCCC 2507

|||||
 Db 48 GGAGAGTGAAGCGGGGTCCAGNCCCTGTGTTGGGCCC 12
 |||||

RESULT 14

BX345406

LOCUS

DEFINITION

BX345406

ACCESSION

BX345406

VERSION

BX345406.1

KEYWORDS

EST.

SOURCE

ORGANISM

Homo sapiens

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

Mammalia; Eutheria; Primates; Catarrhini; Hominoidea; Homo.

Li, W.B., Gruber, C., Jessee, J. and Polayes, D.

Full-length cDNA libraries and normalization

Unpublished (2001)

Contact: Genoscope

Genoscope - Centre National de Sequencage

BP 191 91006 EVRY cedex - France

Email: segref@genoscope.cns.fr, Web : www.genoscope.cns.fr

Library was constructed by Life Technologies, a division of

Invitrogen. This sequence belongs to sequence cluster 3529.f for

more information about this cluster, see

http://www.genoscope.cns.fr/

cgi-bin/cluster.cgi?seq=CS0AC023DA05QPlcluster=3529.f. Contact :

Feng Liang Email : fliang@lifetech.com URL :

http://fulllength.invitrogen.com/ Invitrogen Corporation 1600

Faraday Avenue Genoscope sequence ID : CS0AC023DA05QPl.

FEATURES

source

1. .939

/organism="Homo sapiens"

/mol_type="mRNA"

/db_xref="taxon:9606"

/clone="CS0DC023YB10"

/tissue_type="NEUROBLASTOMA COT 25-NORMALIZED"

/clone_lib="Homo sapiens NEUROBLASTOMA COT 25-NORMALIZED"

/note="1st strand cDNA was primed with a NotI-oligo (dT)

primer. Five prime end enriched, double-strand cDNA was

digested with Not I and cloned into the Not I and EcoR V

sites of the pCMVSPORT 6 vector. Library was normalized."

ORIGIN

Query Match

Best Local Similarity

Matches

755; Conservative

22; Mismatches

20; Indels

6; Gaps

5;

QY 1 AGCTGGGCTCCGGGTGAGGGCTAAAGCGGCTCCCGCCCGGGGCCCCCGCCG 60

Db 142 AGCTGGGCTCCGGGTGAGGGCTAAAGCGGCTCCCGCCCGGGGCCCCCGCCG 201

QY 61 GCCCGCCCGCTGCCCGCCCGCCCGCCCGCCCGCCCGCCCGCCCGCCGTC 120

Db 202 GCCCGCCCGCTGCCCGCCCGCCCGCCCGCCCGCCCGCCCGCCGTC 261

QY 121 TGGGCTATGCTTCGCGGCTTGGCTCGCGGCTCGGGTCCCGAGAGTCCACCGTGG 180

Db 262 TGGGCTATGCTTCGCGGCTTGGCTCGCGGCTCGGGTCCCGAGAGTCCACCGTGG 321

QY 181 CCAACCCAGTCTGCTGTCACCCGACCTGCTTCCCACTTCTCTGTTGGAGCCCGAGG 240

Db 322 CCAACCCAGTCTGCTGTCACCCGACCTGCTTCCCACTTCTCTGTTGGAGCCCGAGG 381

QY 241 ATGTGTACATCGTCAAGAACAAAGCCAGTGTGTGTGTGCAAGGCGCTGCCCGCAGC 300

Db 382 ATGTGTACATCGTCAAGAACAAAGCCAGTGTGTGTGTGCAAGGCGCTGCCCGCAGC 439

QY 301 AGATCTCTTCAAGTCCAAAGGGAGTGGGTGGCGCCAGGTGGACACAGTGTATCGAGCGCA 360

|||||

mis Page Blank (uspio)

GenCore version 5.1.6
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: October 4, 2004, 18:10:30 ; Search time 137 Seconds
(without alignments)
1854.089 Million cell updates/sec

Title: US-09-970-944-2

Perfect score: 4787

Sequence: 1 MAVRGLNPALLGIVLAAML.....AVAGLQPDAGLFTVSEAPC 899

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 1586107 seqs, 282547505 residues

Total number of hits satisfying chosen parameters: 1586107

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : A_Geneseq_29Jan04.*

- 1: Geneseq1980s.*
- 2: Geneseq1990s.*
- 3: Geneseq2000s.*
- 4: Geneseq2001s.*
- 5: Geneseq2002s.*
- 6: Geneseq2003as.*
- 7: Geneseq2003bs.*
- 8: Geneseq2004s.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	4787	100.0	899	AAU79939	Aau79939 Human UNC
2	4688.5	98.2	898	AAU85403	Aau85403 Human pro
3	4688.5	97.9	898	AAU97899	Aau97899 Human net
4	4545.5	95.0	898	AAW78898	Aaw78898 Rat UNC-5
5	4545.5	95.0	898	AAU10543	Aau10543 Rat netri
6	4545.5	95.0	898	AAU97900	Aau97900 Rat netri
7	4434	92.6	943	AAW79128	Aaw79128 Human pro
8	4379.5	91.5	842	AAU74818	Aau74818 Human REP
9	2789	55.8	556	AAW78899	Aaw78899 Human UNC
10	2673.5	55.8	931	AAW50691	Aab50691 Human UNC
11	2673.5	55.8	931	AD663098	Ade663098 Human pro
12	2673.5	55.8	982	ABG11551	Abg11551 Novel hum
13	2497	52.2	945	AD663096	Ade663096 Rat Prote
14	2490.5	52.0	946	ABG61795	Abg61795 Novel UNC
15	2490	52.0	943	AAW78900	Aaw78900 Rat UNC-5
16	2484	51.9	933	AAO18734	Aao18734 Human NOV
17	2484	51.9	933	AAO18735	Aao18735 Human NOV
18	2479	51.8	945	AAU12244	Aau12244 Human PRO
19	2479	51.8	945	ABO17688	Abol17688 Novel hum
20	2479	51.8	945	ABU80942	Abu80942 Human PRO
21	2479	51.8	945	ABU66642	Abu66642 Human PRO
22	2479	51.8	945	ABU59723	Abu59723 Novel sec
23	2479	51.8	945	ABO24913	Abol24913 Human sec
24	2479	51.8	945	ABU66918	Abu66918 Human sec
25	2479	51.8	945	ADA45665	Ada45665 Novel hum

ALIGNMENTS

RESULT 1

AAU79939
ID AAU79939 standard; protein; 899 AA.

XX
AC AAU79939;

XX
DT 15-JUL-2002 (first entry)

XX
DE Human UNC5-like protein NOV1.

XX
KW Human; NOVX polypeptide; cardiomyopathy; atherosclerosis; cancer;
cell signal processing; metabolic pathway modulation; cancerous tissue;
antibody; diabetes; transgenic animal; UNC5-like protein; NOV1;
chromosome 13.

XX
OS Homo sapiens.

XX
PN WO200229038-A2.

XX
PD 11-APR-2002.

XX
PF 04-OCT-2001; 2001WO-US031377.

XX
PR 04-OCT-2000; 2000US-0237862P.

XX
PA (CURA-) CURAGEN CORP.

XX
PI Herrmann JL, Rastelli L, Shimkets RA;

XX
WPI: 2002-340104/37.

XX
N-PSDB; ABK49422.

XX
PT Novel isolated NOVX polypeptide, and encoded polynucleotide, useful for
treating cardiomyopathy, atherosclerosis, and cancer.

XX
PS Claim 1; Page 9; 180pp; English.

XX
CC The present invention relates to a new NOVX polypeptide having a 900
(NOV1), 4349 (NOV2), 940 (NOV3), 798 (NOV4), 865 (NOV5), or 311 (NOV6)
residue amino acid sequence, as given in the specification. The novel
polypeptide, and its encoding polynucleotide, are used to treat
cardiomyopathy, atherosclerosis, cancer or a disease related to cell
signal processing and metabolic pathway modulation, in a human. Detecting
the polypeptide or polynucleotide is useful for identifying cancerous
tissue. The antibody can be used to treat diabetes or cancer. The host
cells can be used to produce non-human transgenic animals useful in drug
screening. The present amino acid sequence is that of the human UNC5-like
protein NOV1 of the invention. This sequence is encoded by the human UNC5

Ada76096 Human PRO
Ada18746 Human PRO
Ada61369 Homo sapi
Adb13154 Novel hum
Adb27695 Human PRO
Ada86174 Novel hum
Adb15738 Human PRO
Ada47524 Human PRO
Ada67319 Human PRO
Adb30326 Human PRO
Ada85622 Novel hum
Ada96834 Human PRO
Ada79138 Human PRO
Adb16479 Human PRO
Ada87277 Novel hum
Ada91571 Novel hum
Adb14634 Human PRO
Adb18595 Novel hum
Ada93810 Human PRO
Adb19706 Novel hum

26 2479 51.8 945 6 ADA76096
27 2479 51.8 945 6 ADA18746
28 2479 51.8 945 6 ADA61369
29 2479 51.8 945 6 ADB13154
30 2479 51.8 945 6 ADB27695
31 2479 51.8 945 6 ADA86174
32 2479 51.8 945 6 ADB15738
33 2479 51.8 945 6 ADA47524
34 2479 51.8 945 6 ADA67319
35 2479 51.8 945 6 ADB30326
36 2479 51.8 945 6 ADA85622
37 2479 51.8 945 6 ADA96834
38 2479 51.8 945 6 ADA79138
39 2479 51.8 945 6 ADB16479
40 2479 51.8 945 6 ADA87277
41 2479 51.8 945 6 ADA91571
42 2479 51.8 945 6 ADB14634
43 2479 51.8 945 6 ADB18595
44 2479 51.8 945 6 ADA93810
45 2479 51.8 945 6 ADB19706

CC	-like NOV1 gene located on chromosome 13	
XX		
SQ	Sequence 899 AA;	
	Query Match 100.0%; Score 4787; DB 5; Length 899;	
	Best Local Similarity 100.0%; Pred. No. 0;	
	Matches 899; Conservative 0; Mismatches	
QY	1 MAVRGLWPALLGIVLAALWLRSGAQASQATVANPVPGANPDLPLPFLVEPDEVIVKPK 60	
Db	1 MAVRGLWPALLGIVLAALWLRSGAQASQATVANPVPGANPDLPLPFLVEPDEVIVKPK 60	
QY	61 VLLVCKAVPATQIFPKNGEWWVRQVDHVIERTSDSSGEPTMEVRINVSROQVEKVGLE 120	
Db	61 VLLVCKAVPATQIFPKNGEWWVRQVDHVIERTSDSSGEPTMEVRINVSROQVEKVGLE 120	
QY	121 EYWCOCVWSSGTTKSKAYIRIARLRNFEQPLAKEVSLQEGIVLPCRPPEGIPPAE 180	
Db	121 EYWCOCVWSSGTTKSKAYIRIARLRNFEQPLAKEVSLQEGIVLPCRPPEGIPPAE 180	
QY	181 VEWLRNEDLVPSLDPNVYITREHSLVVRQARLADTANYTCVAKNIVARRRSASAAVIVY 240	
Db	181 VEWLRNEDLVPSLDPNVYITREHSLVVRQARLADTANYTCVAKNIVARRRSASAAVIVY 240	
QY	241 VNGGWSWTWMSVCSASGRGWQKRSCTNPAPLNGGAFCEGQNVHDTVSSLLVSDG 300	
Db	241 VNGGWSWTWMSVCSASGRGWQKRSCTNPAPLNGGAFCEGQNVHDTVSSLLVSDG 300	
QY	301 SWSPKWSACGLDCTHWSRECSDPAPRNGEECGTDLDTNRCTSDLCVHSASGPEV 360	
Db	301 SWSPKWSACGLDCTHWSRECSDPAPRNGEECGTDLDTNRCTSDLCVHSASGPEV 360	
QY	361 ALYVGLIAVAVCLVLLVLLVYCRKKEGLSDVADSSILTSGFQPVSKSKADNPHL 420	
Db	361 ALYVGLIAVAVCLVLLVLLVYCRKKEGLSDVADSSILTSGFQPVSKSKADNPHL 420	
QY	421 LTIQPOLSTTTTYQGSCLPRQDQSPKQFTNGHLLSLPLGGGRHTLHSSPTSEAEFVS 480	
Db	421 LTIQPOLSTTTTYQGSCLPRQDQSPKQFTNGHLLSLPLGGGRHTLHSSPTSEAEFVS 480	
QY	481 RLSQNYFSLPRLPGTSMYTGTFNFGRLMTPNTGILLSIPDAPLPRGIYIYITLHK 540	
Db	481 RLSQNYFSLPRLPGTSMYTGTFNFGRLMTPNTGILLSIPDAPLPRGIYIYITLHK 540	
QY	541 PEDVRLPLAGCQTLLSPIVSCGPPGVLLTRPVILAMDHCEPSPDSWSLRLKKQSGSW 600	
Db	541 PEDVRLPLAGCQTLLSPIVSCGPPGVLLTRPVILAMDHCEPSPDSWSLRLKKQSGSW 600	
QY	601 EQDVLHLGEEAPSHLYYCQLEASACVFTFQLGRFALVGEALSVAARKLKLLEAPVAC 660	
Db	601 EQDVLHLGEEAPSHLYYCQLEASACVFTFQLGRFALVGEALSVAARKLKLLEAPVAC 660	
QY	661 TSLEYNIRVYCLHDTHDALKEVVOLEKQLGGQLIQPRVLHFKDSYHNLRLSLHDVPSL 720	
Db	661 TSLEYNIRVYCLHDTHDALKEVVOLEKQLGGQLIQPRVLHFKDSYHNLRLSLHDVPSL 720	
QY	721 WSKLLVSVQEIIPFYHIWNGTQRYLHCTFTLERSVSTSDLACKLWQVVEGQGSFSIN 780	
Db	721 WSKLLVSVQEIIPFYHIWNGTQRYLHCTFTLERSVSTSDLACKLWQVVEGQGSFSIN 780	
QY	781 FNITKTQTRFAELALESEAGVPALVGPSAPKIPFLIRQKILSLDPPCRRGADWRTLAQK 840	
Db	781 FNITKTQTRFAELALESEAGVPALVGPSAPKIPFLIRQKILSLDPPCRRGADWRTLAQK 840	
QY	841 LHLDSHLSFSPKSPPTAMILNWEARHFFPNGLSQLAAVAGLQPDAGLFTVSEAC 899	
Db	841 LHLDSHLSFSPKSPPTAMILNWEARHFFPNGLSQLAAVAGLQPDAGLFTVSEAC 899	
XX	AAU85403 standard; protein; 898 AA.	

AC	AAU85403;	
XX		
DT	21-MAY-2002 (first entry)	
XX		
DE	Human protein NOV1.	
XX		
KW	Human; NOVX; cardiomyopathy; atherosclerosis; diabetes;	
KW	cell signal processing disorder; metabolic disorder; obesity; infection;	
KW	anorexia; cancer-associated cachexia; cancer; neurodegenerative disorder;	
KW	Alzheimer's disease; Parkinson's disease; immune disorder;	
KW	haematopoietic disorders; dyslipidaemia; pain; asthma; hypertension;	
KW	osteoporosis; Crohn's disease; multiple sclerosis; angina pectoris;	
KW	myocardial infarction; ulcer; allergy; benign prostatic hypertrophy;	
KW	psychosis; neurological disorder; anxiety; schizophrenia;	
KW	manic depression; dementia; dyskinesia; Huntington's disease;	
KW	Gilles de la Tourette's syndrome; gene therapy.	
XX		
OS	Homo sapiens.	
XX		
FN	WO200210216-A2.	
XX		
PD	07-FEB-2002.	
XX		
PF	30-JUL-2001; 2001WO-US024225.	
XX		
PR	28-JUL-2000; 2000US-0221409P.	
PR	04-AUG-2000; 2000US-0222840P.	
PR	04-AUG-2000; 2000US-0223752P.	
PR	04-AUG-2000; 2000US-0223762P.	
PR	04-AUG-2000; 2000US-0223769P.	
PR	04-AUG-2000; 2000US-0223770P.	
PR	14-AUG-2000; 2000US-0225146P.	
PR	15-AUG-2000; 2000US-0225392P.	
PR	15-AUG-2000; 2000US-0225470P.	
PR	16-AUG-2000; 2000US-0225697P.	
PR	01-FEB-2001; 2001US-0263662P.	
PR	05-APR-2001; 2001US-0281645P.	
XX		
PA	(CURA-) CURAGEN CORP.	
XX		
PI	Padigar M, Mezes P, Mishra V, Burgess C, Casman S, Grose WM;	
PI	Alsobrook JP, Lepley DM, Gerlach VL, Macdougall JR, Smithson G;	
XX		
XX	WPI: 2002-180074/23.	
DR	N-PSDB; ABK37922.	
XX		
PT	New isolated cytoplasmic, nuclear, membrane bound, or secreted	
PT	polypeptide, useful for treating cardiomyopathy, atherosclerosis,	
PT	infections, cancer, neurodegenerative, metabolic, hematopoietic and	
PT	immune disorders.	
XX		
PS	Claim 1; Page 11; 213pp; English.	
XX		
CC	The invention relates to an isolated cytoplasmic, nuclear, membrane	
CC	bound, or secreted polypeptide (NOVX, x= 1-14) their variants or mature	
CC	form. Also included are the nucleic acids encoding the NOVX proteins, a	
CC	vector comprising the nucleic acid, a cell comprising the vector, an anti	
CC	-NOVX antibody and modulators of NOVX. NOVX, the nucleic acid and the	
CC	antibody are useful for treating or preventing a NOVX-associated	
CC	disorder, where the disorder is selected from cardiomyopathy,	
CC	atherosclerosis, diabetes, a disorder related to cell signal processing	
CC	and metabolic pathway modulation, metabolic disorders, obesity,	
CC	infectious disease, anorexia, cancer-associated cachexia, cancer,	
CC	neurodegenerative disorders, Alzheimer's disease, Parkinson's disease,	
CC	immune disorders, haematopoietic disorders, and the various	
CC	dyslipidaemias, metabolic disturbances associated with obesity, the	
CC	metabolic syndrome X and wasting disorders associated with chronic	
CC	diseases, bacterial, fungal, protozoal and viral infections, pain,	
CC	bulimia, asthma, hypertension, urinary retention, osteoporosis, Crohn's	
CC	disease, multiple sclerosis, Albricht Hereditary Osteodystrophy, angina	
CC	pectoris, myocardial infarction, ulcer, allergy, benign prostatic	
CC	hypertrophy, and psychotic and neurological disorders, including anxiety,	
CC	schizophrenia, manic depression, delirium, dementia, and dyskinesias,	
CC		

CC such as Huntington's disease and Gilles de la Tourette's syndrome. The
 CC nucleic acid is useful in gene therapy. The present sequence represents a
 CC NOVX protein
 XX
 SQ Sequence 898 AA;

Query Match 98.2%; Score 4698.5; DB 5; Length 898;
 Best Local Similarity 98.7%; Pred. No. 0;
 Matches 888; Conservative 2; Mismatches 7; Indels 3; Gaps 3;

QY 1 MAVRPGMLPALLGIVLAAMLRGSGAQSATVAVPFGANDPDLPHFLVPEDEVYIVKNKP 60
 DB 1 MAVRPGMLPALLGIVLAAMLRGSGAQSATVAVPFGANDPDLPHFLVPEDEVYIVKNKP 60
 QY 61 VLLVCKAVPATQIFFKCNCEWVRQVDHVIETSDSGSGEPTMEVRINVSROQVEKVGLE 120
 DB 61 VLLVCKAVPATQIFFKCNCEWVRQVDHVIETSDSGSGEPTMEVRINVSROQVEKVGLE 120
 QY 121 EYWCQCVAMSSSGTTKSKAYIRIARLRKNFEQELAKEVSLFOGIVLPCRPEGIPPAE 180
 DB 121 EYWCQCVAMSSSGTTKSKAYIRIARLRKNFEQELAKEVSLFOGIVLPCRPEGIPPAE 180
 QY 181 VEWLRNEDLVDPDLDPNVYITREHSLVVRQARLADTANYTCVAKNIVARRRSASAAVIVY 240
 DB 181 VEWLRNEDLVDPDLDPNVYITREHSLVVRQARLADTANYTCVAKNIVARRRSASAAVIVY 240
 QY 241 VNGGWSTWESVCSASCGRGWKQKRSCTNPAPLNGGAFCEGQNVHDTVSLSLVSDG 300
 DB 241 VNGGWSTWESVCSASCGRGWKQKRSCTNPAPLNGGAFCEGQNVHDTVSLSLVSDG 300
 QY 301 SWSPMKWSACGLDCTHWRSCSDPAPRNGGECQGTDLTRNCTSDLCVHSASGPEV 360
 DB 301 SWSPMKWSACGLDCTHWRSCSDPAPRNGGECQGTDLTRNCTSDLCVHSASGPEV 360
 QY 361 ALYVGLIAVAVCLVLLVLLVLYCRKKEGLSDVADSSILTSQPVSIKPSKADNPHL 420
 DB 361 ALYVGLIAVAVCLVLLVLLVLYCRKKEGLSDVADSSILTSQPVSIKPSKADNPHL 420
 QY 421 LTIQPDLTSTTTTQGSCLCPROGSPKPFQNLNGLHSLPLGGGRHTLHSSPTSABEFV 479
 DB 421 LTIQPDLTSTTTTQGSCLCPROGSPKPFQNLNGLHSLPLGGGRHTLHSSPTSABEFV 479
 QY 480 SRLSTQNYFRSLPRGTSNMTYGTNFGRLMIPNTGISLLIPDIPRGIYEIYVTLH 539
 DB 480 SRLSTQNYFRSLPRGTSNMTYGTNFGRLMIPNTGISLLIPDIPRGIYEIYVTLH 539
 QY 540 KPEDVRLPLAGCOTLLSPVSCGPPVLLTRPVILAMDHCGEPPSDSWSLRKXQSCGS 599
 DB 540 KPEDVRLPLAGCOTLLSPVSCGPPVLLTRPVILAMDHCGEPPSDSWSLRKXQSCGS 599
 QY 600 WEQDVLHGEAPSHLYYQLEASACVYFTQGRFALVGEALSVAARLKLILFAPVA 659
 DB 600 WEQDVLHGEAPSHLYYQLEASACVYFTQGRFALVGEALSVAARLKLILFAPVA 659
 QY 660 CTSLEYNIRVYCLHDTDALKEVQLEKQGGQIQEPRVLHFKDSYHNLRSLSHDVPSS 719
 DB 660 CTSLEYNIRVYCLHDTDALKEVQLEKQGGQIQEPRVLHFKDSYHNLRSLSHDVPSS 719
 QY 720 LWSKLLVSYQEIIPFYHWNCTORYLHCTFTLERSVSPSTSLACKLWVQVGEQGQFSI 779
 DB 720 LWSKLLVSYQEIIPFYHWNCTORYLHCTFTLERSVSPSTSLACKLWVQVGEQGQFSI 779
 QY 780 NFNTTKDTRFAELLALSEAGVPALVGPSAPKIPFLIRQKIISLDPPCRGADWRTIAQ 839
 DB 780 NFNTTKDTRFAELLALSEAGVPALVGPSAPKIPFLIRQKIISLDPPCRGADWRTIAQ 839
 QY 840 KLHLDLSLSPFASKPSPTAMILNLWEARHFPNGLSQLAAVAGLQPDAGLFTVSAEC 899
 DB 840 KLHLDLSLSPFASKPSPTAMILNLWEARHFPNGLSQLAAVAGLQPDAGLFTVSAEC 899

RESULT 3
 AAU97899

ID AC AAU97899 standard; protein; 898 AA.
 XX AAU97899;
 DT 27-AUG-2002 (first entry)
 XX Human netrin binding membrane receptor UNC5H-1 protein.
 KW Netrin binding membrane receptor; receptor; UNC5H-1; human; neurotropic;
 KW neuroprotective; cytosolic; antiparkinsonian; cerebroprotective; cancer;
 KW central nervous system; CNS; stroke; Parkinson's disease;
 KW multiple sclerosis; Alzheimer's disease.
 XX
 OS Homo sapiens.
 XX
 FH Key
 FT Domain 152..223 Location/Qualifiers
 FT Domain /note="Immunoglobulin domain "
 FT Domain 247..294
 FT Domain /note="Thrombospondine type 1 domain "
 FT Domain 302..348
 FT Region /note="Thrombospondine type 1 domain"
 FT Region /note="Transmembrane region"
 FT Domain 495..598
 FT Domain /note="ZUS domain"
 FT Domain 817..897
 FT Domain /note="Death domain"
 XX
 PN WO200233080-A2.
 XX
 PD 25-APR-2002.
 XX
 PF 15-OCT-2001; 2001WO-EP011891.
 XX
 PR 16-OCT-2000; 2000US-0240061P.
 XX
 PA (PAB) BAYER AG.
 XX
 PI Koehler RH;
 XX
 DR WPI: 2002-463314/49.
 XX N-PSDB; ABKS2891.
 XX
 PT Novel human netrin binding membrane receptor polypeptide and
 PT polynucleotides for identifying modulating agents useful in treating
 PT diseases e.g. Parkinson's disease, multiple sclerosis, stroke,
 XX Alzheimer's disease.
 PS Claim 1; Fig 2; 94pp; English.
 XX
 CC This invention relates to the DNA and protein sequences of a novel
 CC purified human netrin binding membrane receptor, UNC5H-1. The DNA
 CC sequence of the invention is useful as a probe for detecting a nucleic
 CC acid encoding the UNC5H-1 protein in a biological sample. The sequences
 CC of the invention are useful to screen for agents which decrease the
 CC activity of the UNC5H-1 protein. The sequences are also useful for
 CC screening agents which regulate (modulate) the activity of the protein of
 CC the invention. A pharmaceutical composition containing the protein of the
 CC invention or a reagent that modulates the activity of the UNC5H-1 protein
 CC may be useful for treating a UNC5H-1 dysfunction related disease such as
 CC cancer or a central nervous system (CNS) disorders (e.g. Parkinson's
 CC disease, multiple sclerosis, stroke and Alzheimer's disease). Fusion
 CC proteins comprising the UNC5H-1 protein are useful for generating
 CC antibodies and for in various assay systems, and the protein can be used
 CC as a bait protein in a two-hybrid assay or three-hybrid assay. The method
 CC of the invention is useful for detecting a coding sequence for the UNC5H-
 CC 1 protein. The present sequence represents the human netrin binding
 CC membrane receptor UNC5H-1 protein of the invention
 XX Sequence 898 AA;

Query Match 97.9%; Score 4688.5; DB 5; Length 898;

Best Local Similarity 98.4%; Pred. No. 0;		Matches 886; Conservative 3; Mismatches 8; Indels 3; Gaps 3;	
QY	1	MAVRPGLWPALLGIVLAALRGSGAQOQATVANPVGANPDLLPHFLVEPEDVYVKNKP	60
Db	1	MAVRPGLWPALLGIVLAALRGSGAQOQATVANPVGANPDLLPHFLVEPEDVYVKNKP	60
QY	61	VLLVCXAVPATQIFPKNGEWRQVDHVIERTDSSGSEPTMEVNRINVSQQVEKVPGLE	120
Db	61	VLLVCXAVPATQIFPKNGEWRQVDHVIERTDSSGSEPTMEVNRINVSQQVEKVPGLE	120
QY	121	EYWCQCVAWSSGTTKSKAYIRIARLNKFEQPLAKEVSLQGIIVLPCRPPEGIPPAE	180
Db	121	EYWCQCVAWSSGTTKSKAYIRIARLNKFEQPLAKEVSLQGIIVLPCRPPEGIPPAE	180
QY	181	VEWLRNEDLVDPSPDPNVYITREHSLVVRQARLADTANYTCVAKNIVARRRSASAAVIVY	240
Db	181	VEWLRNEDLVDPSPDPNVYITREHSLVVRQARLADTANYTCVAKNIVARRRSASAAVIVY	240
QY	241	VNGGWSTWTEWVCSASCGRGHQKRSRSCNTPAPLNGGAFCEGQNV-QKTACATLCPVDG	300
Db	241	VNGGWSTWTEWVCSASCGRGHQKRSRSCNTPAPLNGGAFCEGQNV-QKTACATLCPVDG	299
QY	301	SWSPKSKWSACGLDCTHWRSCSDPAPNGGEECGTDLTRNCTSDLCVHSASGPEV	360
Db	300	SWSPKSKWSACGLDCTHWRSCSDPAPNGGEECGTDLTRNCTSDLCVHSASGPEV	359
QY	361	ALYVGLIAVAVCLVLLVLLVYCRKKEGLSDVADSSILTSFGQFVSIKPSKADNPHL	420
Db	360	ALYVGLIAVAVCLVLLVLLVYCRKKEGLSDVADSSILTSFGQFVSIKPSKADNPHL	419
QY	421	LTIQPDLS-TTTTYQSLCPRQDGPSPKQTLNGLHLLSPGGRHTLHSSSTSEAEV	479
Db	420	LTIQPDLS-TTTTYQSLCPRQDGPSPKQTLNGLHLLSPGGRHTLHSSSTSEAEV	479
QY	480	SRLSTQNYFRSLPRGTSNMTYGFNFGRLMIPNTGISILLIPDPAIPRGKIYIYLTILH	539
Db	480	SRLSTQNYFRSLPRGTSNMTYGFNFGRLMIPNTGISILLIPDPAIPRGKIYIYLTILH	539
QY	540	KPEDVRLPLAGCQTLLSPVSCGPPGVLLTRPVILAMDHCGEPPSPDSWSLRLKKQCEGS	599
Db	540	KPEDVRLPLAGCQTLLSPVSCGPPGVLLTRPVILAMDHCGEPPSPDSWSLRLKKQCEGS	599
QY	600	WEQDVLHGEAPSHLYYCOLEASACVVFTEQLGRFALVGEALSVAAAKELKLLFAPVA	659
Db	600	WE-DVLHGEAPSHLYYCOLEASACVVFTEQLGRFALVGEALSVAAAKELKLLFAPVA	658
QY	660	CTSEYNIIRVYCLHDTHDALKEVVQLEKQGGQIOEPVILHFKDSYHNLRLSIHDVPSS	719
Db	659	CTSEYNIIRVYCLHDTHDALKEVVQLEKQGGQIOEPVILHFKDSYHNLRLSIHDVPSS	718
QY	720	LWKSLLVSYQEIIPFYHNGTQRYLHCTFTLERSVSPSTSDLACKLWVWQVGGQSF	779
Db	719	LWKSLLVSYQEIIPFYHNGTQRYLHCTFTLERSVSPSTSDLACKLWVWQVGGQSF	778
QY	780	NFNITKDTREAEILLAESEAGVPAVCSAFKIPFLIRQKIIISLDPPCCRGADWRTLAQ	839
Db	779	NFNITKDTREAEILLAESEAGVPAVCSAFKIPFLIRQKIIISLDPPCCRGADWRTLAQ	838
QY	840	KLHLDLSLSPFASKPSPTAMLLNWEARHPNGNLSQALAAVAGLQPDAGLFTVSEAC	899
Db	839	KLHLDLSLSPFASKPSPTAMLLNWEARHPNGNLSQALAAVAGLQPDAGLFTVSEAC	898
RESULT 4			
ID	AAW78998 standard; protein; 898 AA.		
XX	AAW78998;		
AC	(revised)		
DT	25-MAR-2003 (first entry)		
DT	21-DEC-1998		
XX			

DE	Rat UNC-5 homologue UNC5H-1.	
XX	UNC-5; UNC5H-1; rat; netrin receptor; cell migration; axon guidance;	
KW	diagnosis; therapy.	
OS	Rattus sp.	
XX	Key Location/Qualifiers	
FT	Peptide 580..594	
XX	/note= "peptide used to raise rabbit polyclonal antiser"	
FN	WO9837085-A1.	
XX	27-AUG-1998.	
XX	19-FEB-1998; 98WO-US003143.	
XX	19-FEB-1997; 97US-00808982.	
XX	(REGC) UNIV CALIFORNIA.	
PA	Tessier-Lavigne M, Leonardo ED, Hinck L, Masu M, Keinomasa K;	
PI	WPI; 1998-495364/42.	
DR	N-PSDB; AAV52940.	
XX	Netrin-binding, vertebrate proteins - useful for diagnosis, therapy and	
PT	the biopharmaceutical industry.	
XX	Claim 1; Page 19-22; 32pp; English.	
XX	UNC5H-1 and UNC5H-2 (see AAW78900) are rat homologues of Caenorhabditis	
CC	elegans UNC-5 protein. Their amino acid sequences were deduced from	
CC	isolated unc5h cDNA clones (see AAV52940 and AAV52942) isolated from an	
CC	E18 brain cDNA library. The predicted proteins show similarity with UNC-	
CC	5, possess 2 predicted Ig-like domains and 2 predicted thrombospondin	
CC	type-1 repeats, a predicted membrane spanning region, and a large	
CC	intracellular domain. They are predicted to be involved in cell migration	
CC	and axon guidance, and are characterized as receptor proteins for	
CC	netrins. Human UNC5H-1 (see AAW78899) and UNC5H-2 (see AAW78901) proteins	
CC	are also claimed. Vertebrate UNC-5 proteins may be produced recombinantly	
CC	from transfected host cells. The invention also provides unc-5	
CC	hybridization probes and primers, vertebrate UNC-5-specific binding	
CC	agents such as specific antibodies, and methods of making and using the	
CC	subject compositions in diagnosis (e.g. genetic hybridisation screens for	
CC	vertebrate unc-5 transcripts), therapy (e.g. gene therapy to modulate	
CC	vertebrate unc-5 gene expression) and in the biopharmaceutical industry	
CC	(e.g. as immunogens, reagents for modulating cell guidance, reagents for	
CC	screening chemical libraries for lead pharmacological agents, etc.).	
CC	(Updated on 25-MAR-2003 to correct PI field.)	
XX	Sequence 898 AA;	
QY	Query Match 95.0%; Score 4545.5; DB 2; Length 898;	
Db	Best Local Similarity 94.7%; Pred. No. 0;	
XX	Matches 852; Conservative 19; Mismatches 26; Indels 3; Gaps 3;	
QY	1	MAVRPGLWPALLGIVLAALRGSGAQOQATVANPVGANPDLLPHFLVEPEDVYVKNKP 60
Db	1	MAVRPGLWPALLGIVLAALRGSGAQOQATVANPVGANPDLLPHFLVEPEDVYVKNKP 60
QY	61	VLLVCXAVPATQIFPKNGEWRQVDHVIERTDSSGSEPTMEVNRINVSQQVEKVPGLE 120
Db	61	VLLVCXAVPATQIFPKNGEWRQVDHVIERTDSSGSEPTMEVNRINVSQQVEKVPGLE 120
QY	121	EYWCQCVAWSSGTTKSKAYIRIARLNKFEQPLAKEVSLQGIIVLPCRPPEGIPPAE 180
Db	121	EYWCQCVAWSSGTTKSKAYIRIARLNKFEQPLAKEVSLQGIIVLPCRPPEGIPPAE 180
QY	181	VEWLRNEDLVDPSPDPNVYITREHSLVVRQARLADTANYTCVAKNIVARRRSASAAVIVY 240
Db	181	VEWLRNEDLVDPSPDPNVYITREHSLVVRQARLADTANYTCVAKNIVARRRSASAAVIVY 240

QY	241	V	N	G	G	S	T	W	T	W	S	V	C	S	A	C	G	R	G	W	Q	K	R	S	R	S	C	T	N	P	A	P	L	N	G	A	F	C	E	Q	N	V	H	D	R	T	V	S	L	L	V	D	G	300						
Db	241	V	N	G	G	S	T	W	T	W	S	V	C	S	A	C	G	R	G	W	Q	K	R	S	R	S	C	T	N	P	A	P	L	N	G	A	F	C	E	Q	N	V	H	D	R	T	V	S	L	L	V	D	G	299						
QY	301	S	N	S	P	W	K	S	K	A	C	G	L	D	C	T	H	W	S	R	E	C	S	D	P	A	P	R	N	G	B	E	C	O	G	T	O	L	D	T	R	N	C	T	S	D	I	C	V	H	S	A	S	G	P	E	D	V	360	
Db	300	S	N	S	P	W	K	S	K	A	C	G	L	D	C	T	H	W	S	R	E	C	S	D	P	A	P	R	N	G	B	E	C	R	G	A	D	L	T	R	N	C	T	S	D	I	C	L	T	A	S	C	P	E	D	V	359			
QY	361	A	L	Y	V	G	L	I	A	V	A	V	C	L	V	L	L	L	L	V	C	R	K	E	G	L	D	S	D	V	A	D	S	S	I	L	T	S	G	F	O	P	V	S	I	K	S	K	A	D	N	P	H	L	420					
Db	360	A	L	Y	I	G	L	V	A	V	A	V	C	L	F	L	L	L	A	L	G	L	I	C	R	K	E	G	L	D	S	D	V	A	D	S	S	I	L	T	S	G	F	O	P	V	S	I	K	S	K	A	D	N	P	H	L	419		
QY	421	L	I	I	Q	P	D	L	S	-	T	T	T	T	T	T	T	T	T	T	T	T	T	T	T	T	T	T	T	T	T	T	T	T	T	T	T	T	T	T	T	T	T	T	T	T	T	T	T	479										
Db	420	L	I	I	Q	P	D	L	S	-	T	T	T	T	T	T	T	T	T	T	T	T	T	T	T	T	T	T	T	T	T	T	T	T	T	T	T	T	T	T	T	T	T	T	T	T	T	479												
QY	480	S	R	L	S	T	Q	N	F	R	S	L	P	R	G	T	S	N	M	T	G	T	F	N	F	L	G	R	L	M	P	N	T	G	I	S	L	L	I	P	D	A	I	P	R	G	K	I	E	Y	L	T	L	H	539					
Db	480	S	R	L	S	T	Q	N	F	R	S	L	P	R	G	T	S	N	M	T	G	T	F	N	F	L	G	R	L	M	P	N	T	G	I	S	L	L	I	P	D	A	I	P	R	G	K	I	E	Y	L	T	L	H	539					
QY	540	K	P	E	D	V	R	L	P	L	A	G	C	O	T	L	L	S	P	T	V	S	C	G	P	G	V	L	L	T	R	P	V	I	L	A	M	D	H	C	G	E	P	S	P	D	S	S	L	R	L	K	K	Q	S	C	E	G	S	599
Db	540	K	P	E	D	V	R	L	P	L	A	G	C	O	T	L	L	S	P	T	V	S	C	G	P	G	V	L	L	T	R	P	V	I	L	A	M	D	H	C	G	E	P	S	P	D	S	S	L	R	L	K	K	Q	S	C	E	G	S	599
QY	600	W	E	Q	D	V	L	H	G	E	E	A	P	S	H	L	Y	C	O	L	E	A	S	A	C	Y	F	T	E	Q	L	G	R	F	A	L	V	G	E	A	L	S	V	A	A	A	K	L	L	L	F	A	P	V	A	659				
Db	600	W	E	-	D	V	L	H	G	E	E	S	P	S	H	L	Y	C	O	L	E	A	G	A	C	Y	F	T	E	Q	L	G	R	F	A	L	V	G	E	A																				

AAU74818	
ID	AAU74818 standard; protein; 842 AA.
XX	
AC	AAU74818;
XX	
DT	23-APR-2002 (first entry)
XX	
DE	Human REPTR 1 protein.
XX	
KW	REPTR: human; antiinflammatory; cytostatic; immunosuppressive; antiviral;
KW	anti-HIV; antiarthritic; anticonvulsant; nootropic; neuroprotective;
KW	antiallergic; antibody; immunogen; endometriosis;
KW	gastrointestinal disorder; gastritis; oesophageal carcinoma;
KW	Crohn's disease; irritable bowel syndrome; ulcerative colitis;
KW	endocrine disease; hypothalamus disorder; Kallman's disease;
KW	autoimmune disease; inflammatory disease; infertility; receptor;
KW	acquired immune deficiency syndrome; AIDS; rheumatoid arthritis; allergy;
KW	osteoarthritis; diabetes mellitus; multiple sclerosis;
KW	systemic lupus erythematosus; cell proliferative disorder; cancer;
KW	developmental disorder; Duchenne muscular dystrophy;
KW	Becker muscular dystrophy; neurological disorder; epilepsy;
KW	Alzheimer's disease; Huntington's disease; reproductive disorder.
XX	
OS	Homo sapiens.
XX	
PN	WO2001198354-A2.
XX	
PD	27-DEC-2001.
XX	
PF	21-JUN-2001; 2001WO-US019942.
XX	
PR	21-JUN-2000; 2000US-0214027P.
PR	25-AUG-2000; 2000US-0228045P.
PR	12-DEC-2000; 2000US-0255104P.
XX	
PA	(INCY) INCYTE GENOMICS INC.
XX	
PI	Griffin JA, Kallick DA, Tribouley CM, Yue H, Nguyen DB, Tang YT;
PI	Lai P, Policky JL, Azimzai Y, Lu DAM, Graul R, Yao MG, Burford N;
PI	Hafalia AJA, Baughn MR, Bhandan O, Patterson C, Yang J, Xu Y;
PI	Gandhi AR, Warren BA, Ding L, Sanjanwala MS, Duggan BM, Lu Y;
XX	
WPI	2002-090432/12.
DR	N-PSDB; ABK15169.
XX	
PT	Twelve human receptors (referred to as REPTR-1 to REPTR-12), useful in
PT	the diagnosis, treatment and prevention of gastrointestinal (e.g.
PT	gastritis), autoimmune/inflammatory (e.g. osteoarthritis) and cell
PT	proliferative (e.g. cancer) disorders.
XX	
PS	Claim 45; Page 111-113; 157pp; English.
XX	
CC	This invention relates to twelve human receptors cDNA sequences referred
CC	to as REPTR-1 to REPTR-12, and the proteins encoded thereby. The
CC	proteins of the invention may have antiinflammatory, cytostatic,
CC	immunosuppressive, antiviral, anti-HIV, antiarthritic, muscular active
CC	general, anticonvulsant, nootropic, neuroprotective, antiallergic
CC	activities. The sequences of the invention may be used to produce REPTR
CC	agonists or antagonists, and the protein sequences may be used to raise
CC	anti-REPTR antibodies. These molecules and the REPTR polynucleotides and
CC	polypeptides of the invention are useful in the diagnosis, treatment and
CC	prevention of gastrointestinal (e.g. gastritis, oesophageal carcinoma,
CC	Crohn's disease, irritable bowel syndrome, ulcerative colitis), endocrine
CC	(e.g. hypothalamus disorder, Kallman's disease), autoimmune/ inflammatory
CC	(e.g. acquired immune deficiency syndrome (AIDS), rheumatoid arthritis,
CC	allergies, osteoarthritis, diabetes mellitus, multiple sclerosis,
CC	systemic lupus erythematosus), cell proliferative (e.g. cancer),
CC	developmental (e.g. Duchenne and Becker muscular dystrophy), neurological
CC	(e.g. epilepsy, Alzheimer's disease, Huntington's disease) and
CC	reproductive (e.g. infertility, endometriosis) disorders. Numerous other
CC	examples of each disorder are given in the specification. The present
CC	sequence represents the human REPTR1 protein sequence of the invention
XX	

SQ	Sequence 842 AA;
Query Match	91.5%; Score 4379.5; DB 5; Length 842;
Best Local Similarity	93.1%; Pred. No. 0;
Matches 838; Conservative	1; Mismatches 2; Indels 59; Gaps 3;
QY	1 MAVRPLGWPALLGIVLAALRSGAQQSATVANPVGANPDLLPHFLVEPDDVYIVKNKP 60
DB	1 MAVRPLGWPALLGIVLAALRSGAQQSATVANPVGANPDLLPHFLVEPDDVYIVKNKP 60
QY	61 VLLVCVAVPATQIIFPKCNQGEWVRQVDHVIERTDSSGCEPTMEVRINVSQQQVEKVGLE 120
DB	61 VLLVCVAVPATQIIFPKCNQGEWVRQVDHVIERTDSSGCEPTMEVRINVSQQQVEKVGLE 120
QY	121 EYWCOCVAVSSGTTKSKAVIRIARLNKFEQELAKEVSLQGIQVILPCRPPEGIPPAE 180
DB	121 EYWCOCVAVSSGTTKSKAVIRIARLNKFEQELAKEVSLQGIQVILPCRPPEGIPPAE 180
QY	181 VEWLNRNEDLVDPDLDPNVYITREHSLVVRQARLADTANYTCVAKNIVARRRSASAIVY 240
DB	181 VEWLNRNEDLVDPDLDPNVYITREHSLVVRQARLADTANYTCVAKNIVARRRSASAIVY 240
QY	241 VNGGWSWTWESVCSACGRGWQKRSCTNPAPLNGGAFCEGQNVHRTVSSLLVSDG 300
DB	241 -----VDG 243
QY	301 SWSPWSKWSACGLDCTHWRSECSDPAPRNGGEECGQTDLTDRNCTSDLCVHSASGPDV 360
DB	244 SWSPWSKWSACGLDCTHWRSECSDPAPRNGGEECGQTDLTDRNCTSDLCVHTASGPDV 303
QY	361 ALYVGLIIVAVCLVLLVLLVILVYCRKKEGLSDVADSSILTSFGQFQVSKSKADNPHL 420
DB	304 ALYVGLIIVAVCLVLLVLLVILVYCRKKEGLSDVADSSILTSFGQFQVSKSKADNPHL 363
QY	421 LTIQPDLS-TTTTYQGLCPRODQSPKQFQNLNGHLISPLGGRRHTLHSSPTSEAEFV 479
DB	364 LTIQPDLS-TTTTYQGLCPRODQSPKQFQNLNGHLISPLGGRRHTLHSSPTSEAEFV 423
QY	480 SRLSTQNYPRSLPRGTSNMVTGTFNPLGRLMIPNTGISLLIPPDAPRGIYEIYLTJH 539
DB	424 SRLSTQNYPRSLPRGTSNMVTGTFNPLGRLMIPNTGISLLIPPDAPRGIYEIYLTJH 483
QY	540 KPEDVRLPLAGCQTLSPISVSCGPGVLLTRPVILAMDHCGEPPSPDSWSLRKQSCGFS 599
DB	484 KPEDVRLPLAGCQTLSPISVSCGPGVLLTRPVILAMDHCGEPPSPDSWSLRKQSCGFS 543
QY	600 WEQDVHLHGBEAPSHLYYQLEASACYVTEQIGREALVGEALSVAAAKELKLLFAPVA 659
DB	544 WE-DVHLHGBEAPSHLYYQLEASACYVTEQIGREALVGEALSVAAAKELKLLFAPVA 602
QY	660 CTSLEYNIRVYCLHDTHDALKKEVQLEKOLGGOLIQEPRVLHFKDSYHNLRLSIHDVPS 719
DB	603 CTSLEYNIRVYCLHDTHDALKKEVQLEKOLGGOLIQEPRVLHFKDSYHNLRLSIHDVPS 662
QY	720 LMKSKLLVSYQEIPIFYHINWGTORYLHCTFTLRRVSPSTSDACKLWVQVEGDSQFSI 779
DB	663 LMKSKLLVSYQEIPIFYHINWGTORYLHCTFTLRRVSPSTSDACKLWVQVEGDSQFSI 722
QY	780 NFNITKDTFAELALAESEAGVPALVGPSAFKIPFLIRQKLIISLDPCCRRGADWTJLAQ 839
DB	723 NFNITKDTFAELALAESEAGVPALVGPSAFKIPFLIRQKLIISLDPCCRRGADWTJLAQ 782
QY	840 KLHLDHSLSPFASKPSPTAMINLWEARHFPNGNLSQLAAVAAGLQDPDAGLFTVSEABC 899
DB	783 KLHLDHSLSPFASKPSPTAMINLWEARHFPNGNLSQLAAVAAGLQDPDAGLFTVSEABC 842
RESULT 9	
AAW78899	
ID	AAW78899 standard; protein; 556 AA.
XX	
AC	AAW78899;
XX	

Best Local Similarity 52.1%; Pred. No. 9.7e-207;		Matches 496; Conservative 150; Mismatches 247; Indels 59; Gaps 18;	
QY	1	MAVRPGLWPA	18;
Db	1	NGARSGARGALLALLCWDRLSQAGKRKSGEVL	60
QY	57	KNKPVLLVCKAVPATQIFKCKNGEWVQVDHVI	111
Db	61	KNKPVLLVCKAVPATQIFKCKNGEWVQVDHVI	120
QY	112	QVEKVFGLLEWYCCVAVSSSTKSKAYIRIARL	171
Db	121	QVEELFGLLEWYCCVAVSSSTKSKAYIRIARL	180
QY	172	PPEGIPPAEVEWLRNEDLVPSLDPNVYITREH	231
Db	181	PPEGIPPAEVEWLRNEDLVPSLDPNVYITREH	240
QY	232	SASAAVIVYVNGWSTWTEWVSCSACGRGWKR	291
Db	241	STTATVIVYVNGWSTWTEWVSCSACGRGWKR	299
QY	292	SSLLSVDSGSPWKSACGLDCTHRSRECS	351
Db	300	CTTICFVDGATEWKSACGLDCTHRSRECS	359
QY	352	HS-----ASGPDVALVGL-IAVAVCLVLL	401
Db	360	QSEPVPAVLEASG--DAALYAGLVAVI	417
QY	402	TSGFOPVSKPANDPHLL--TIQPDLSITT	458
Db	418	TSGFOPVSKPANDPHLL--TIQPDLSITT	476
QY	459	LGGRHTLHHSPT-----SAAEFVSLSTQNY	490
Db	477	LPSSLKVVYSSTTSGGGLAGADLLGVLP	536
QY	491	LPRGTSNMTYGFNLGGRMIFNTGISILL	549
Db	537	LPRDGGSSVSGTFCGLGGRLSIPGTG	595
QY	550	GCOTLLSPITVSGCPGVLLTRPVILAMH	609
Db	596	GTQTVLSPSVTGCPTGLLCRPVILTMPH	655
QY	610	BAPSHLYYQLEASACVVFTEQLGREAL	659
Db	656	ETLNTFCYQLEPRACHILLDQLGT	715
QY	670	YCLHDTHDALKEVYQLEKQGLGQI	729
Db	716	YCLEDTFVALKEVLELRTLGGVLEEP	775
QY	730	QETPFYHNGTORYLHCTFTLREVSPTS	789
Db	776	QETPFYHNGTORYLHCTFTLREVSPTS	834
QY	790	AEILLALESEAG--VPALVGPSAFKIP	847
Db	835	GSLDTCSPAGSVITQLGPAFKIP	894
QY	848	SPTFASKPSPTAMILNWEARHFNPN	899
Db	895	NYFATKASPTGVILDLWEALQDDG	946
RESULT 15			
AAW78900			
ID AAW78900 standard; protein; 943 AA.			
XX			
AC AAW78900;			

XX	25-MAR-2003	(revised)	
DT	21-DEC-1998	(first entry)	
XX			
DE			
XX			
KW	Rat UNC-5 homologue UNC5H-2.		
XX	UNC-5; UNC5H-2; rat; netrin receptor; cell migration; axon guidance; diagnosis; therapy.		
OS	Rattus sp.		
XX			
Key	Location/Qualifiers		
FT	Peptide	148..161	
FT	/note= "peptide used to raise rabbit polyclonal antisera"		
FT	Misc-difference	753	
FT	/note= "encoded by CG"		
FT	Peptide	909..924	
FT	/note= "peptide used to raise rabbit polyclonal antisera"		
XX			
PN	W09837085-A1.		
XX			
PD	27-AUG-1998.		
XX			
PF	19-FEB-1998; 98WO-US003143.		
XX			
PR	19-FEB-1997; 97US-00808982.		
XX			
PA	(REGC) UNIV CALIFORNIA.		
XX			
PI	Tessier-Lavigne M, Leonardo ED, Hinck L, Masu M, Keinomasu K;		
XX			
DR	WPI; 1998-495364/42.		
DR	N-PSDB; AAV52942.		
XX			
PT	Netrin-binding, vertebrate proteins - useful for diagnosis, therapy and the biopharmaceutical industry.		
XX			
PS	Claim 1; Page 24-26; 32pp; English.		
XX			
CC	UNC5H-1 and UNC5H-2 (see AAW78900) are rat homologues of Caenorhabditis elegans UNC-5 protein. Their amino acid sequences were deduced from an isolated unc5H cDNA clones (see AAV52940 and AAV52942) isolated from an E18 brain cDNA library. The predicted proteins show similarity with UNC-5, possess 2 predicted Ig-like domains and 2 predicted thrombospondin type-1 repeats, a predicted membrane spanning region, and a large intracellular domain. They are predicted to be involved in cell migration and axon guidance, and are characterised as receptor proteins for netrins. Human UNC5H-1 (see AAW78899) and UNC5H-2 (see AAW78901) proteins are also claimed. Vertebrate UNC-5 proteins may be produced recombinantly from transfected host cells. The invention also provides unc-5 hybridisation probes and primers. The vertebrate UNC-5-specific binding agents such as specific antibodies, and methods of making and using the subject compositions in diagnosis (e.g. genetic hybridisation screens for vertebrate unc-5 transcripts), therapy (e.g. gene therapy to modulate vertebrate unc-5 gene expression) and in the biopharmaceutical industry (e.g. as immunogens, reagents for modulating cell guidance, reagents for screening chemical libraries for lead pharmacological agents, etc.).		
CC	(Updated on 25-MAR-2003 to correct PI field.)		
XX			
Sequence 943 AA;			
Query Match	52.0%; Score 2490; DB 2; Length 943;		
Best Local Similarity	52.3%; Pred. No. 1.1e-206;		
Matches	496; Conservative 144; Mismatches 226; Indels 82; Gaps 19;		
QY	9	PALLGIVLAARLGRSGAQQSATVANPVGANPDL	68
Db	21	PSLAGI-----DSGAQ---GLPDSFP	70
QY	69	PATQIPKCKNGEWVQVDHVIERTDSSGSEPT	128
Db	71	PAIQIVKCKNGEWVSKGHVTOESLDEATG	130

This Page Blank (uspto)

GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - protein search, using sw model

Run on: October 4, 2004, 18:26:37 ; Search time 34 seconds
(without alignments)
1365.051 Million cell updates/sec

Title: US-09-970-944-2
Perfect score: 4787
Sequence: 1 MAVRGLWPALLGIVLAAML.....AVAGLQPDAGLFTVSEABC 899

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 389414 seqs, 51625971 residues

Total number of hits satisfying chosen parameters: 389414

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : Issued Patents AA:
1: /cgn2_6/ptodata/2/1aa/5A_COMB.pep:*
2: /cgn2_6/ptodata/2/1aa/5B_COMB.pep:*
3: /cgn2_6/ptodata/2/1aa/6A_COMB.pep:*
4: /cgn2_6/ptodata/2/1aa/6B_COMB.pep:*
5: /cgn2_6/ptodata/2/1aa/PCTUS_COMB.pep:*
6: /cgn2_6/ptodata/2/1aa/backfiles1.pep:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Match	Length	ID	Description
1	4545.5	95.0	898	US-08-808-982-5	Sequence 5, Appli
2	4545.5	95.0	898	US-09-306-902A-5	Sequence 5, Appli
3	2789.5	58.3	557	US-08-808-982-6	Sequence 6, Appli
4	2789.5	58.3	557	US-09-306-902A-6	Sequence 6, Appli
5	2490.5	52.0	943	US-08-808-982-7	Sequence 7, Appli
6	2490.5	52.0	943	US-09-306-902A-7	Sequence 7, Appli
7	294.6	6.1	102	US-08-808-982-8	Sequence 8, Appli
8	294.6	6.1	102	US-09-306-902A-8	Sequence 8, Appli
9	273.5	5.7	1172	US-08-313-288B-19	Sequence 19, Appli
10	249.5	5.2	441	US-08-985-526-3	Sequence 3, Appli
11	246.5	5.1	239	PCT-US93-01652-1	Sequence 1, Appli
12	246.5	5.1	1170	US-08-313-288B-20	Sequence 20, Appli
13	235.5	4.9	218	US-08-985-526-1	Sequence 1, Appli
14	234.5	4.9	1395	US-09-540-245A-15	Sequence 15, Appli
15	216.5	4.5	1651	US-09-540-245A-18	Sequence 18, Appli
16	208.5	4.4	469	US-08-313-288B-15	Sequence 15, Appli
17	205.5	4.3	1381	US-09-540-245A-16	Sequence 16, Appli
18	198.5	4.1	788	US-08-918-914-4	Sequence 4, Appli
19	194.5	4.1	1069	US-08-877-730-2	Sequence 2, Appli
20	194.5	4.1	1150	US-09-877-730-8	Sequence 8, Appli
21	189.5	3.9	904	US-09-877-730-6	Sequence 6, Appli
22	189.5	3.9	985	US-09-877-730-10	Sequence 10, Appli
23	188.5	3.9	1297	US-09-540-245A-17	Sequence 17, Appli
24	186.5	3.9	380	US-09-877-730-4	Sequence 4, Appli
25	186.5	3.9	1266	US-08-506-296B-4	Sequence 4, Appli
26	168.5	3.5	481	US-09-130-491-8	Sequence 8, Appli
27	167.5	3.5	905	US-09-369-364A-9	Sequence 9, Appli

28	166	3.5	837	4	US-09-122-126B-2	Sequence 2, Appli
29	166	3.5	837	4	US-09-634-286A-2	Sequence 2, Appli
30	160.5	3.4	949	4	US-09-568-559-2	Sequence 2, Appli
31	160.5	3.4	967	4	US-09-130-491-2	Sequence 2, Appli
32	159.5	3.3	727	4	US-09-445-023A-1	Sequence 1, Appli
33	159.5	3.3	1224	4	US-09-930-872-4	Sequence 4, Appli
34	158	3.3	608	4	US-09-130-491-13	Sequence 13, Appli
35	156	3.3	997	4	US-09-369-364A-7	Sequence 7, Appli
36	153.5	3.2	1059	4	US-09-907-794A-290	Sequence 290, App
37	153.5	3.2	1059	4	US-09-905-125A-290	Sequence 290, App
38	153.5	3.2	1059	4	US-09-902-775A-290	Sequence 294, App
39	153.5	3.2	1119	4	US-09-907-794A-294	Sequence 294, App
40	153.5	3.2	1119	4	US-09-905-125A-294	Sequence 294, App
41	153.5	3.2	1119	4	US-09-902-775A-294	Sequence 17, Appli
42	153	3.2	1081	4	US-09-369-364A-17	Sequence 17, Appli
43	152.5	3.2	757	4	US-09-963-791-24	Sequence 24, Appli
44	152.5	3.2	908	4	US-09-963-791-2	Sequence 2, Appli
45	152.5	3.2	1911	1	US-08-348-008B-5	Sequence 5, Appli

ALIGNMENTS

RESULT 1
US-08-808-982-5
; Sequence 5, Application US/08808982
; Patent No. 5939271
; GENERAL INFORMATION:
; APPLICANT: Tessier-Lavigne, Marc
; APPLICANT: Leonardo, E. David
; APPLICANT: Hink, Lindsay
; APPLICANT: Masu, Masayuki
; APPLICANT: Kazuko, Keino-Masu
; TITLE OF INVENTION: Netrin Receptors
; NUMBER OF SEQUENCES: 8
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: SCIENCE & TECHNOLOGY LAW GROUP
; STREET: 268 BUSH STREET, SUITE 3200
; CITY: SAN FRANCISCO
; STATE: CALIFORNIA
; COUNTRY: USA
; ZIP: 94104
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/808,982
; FILING DATE:
; CLASSIFICATION: 530
; ATTORNEY/AGENT INFORMATION:
; NAME: OSMAN, RICHARD A
; REGISTRATION NUMBER: 36,627
; REFERENCE/DOCKET NUMBER: UC96-217
; TELEPHONE: (415) 343-4341
; TELEFAX: (415) 343-4342
; INFORMATION FOR SEQ ID NO: 5:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 898 amino acids
; TYPE: amino acid
; STRANDEDNESS: not relevant
; TOPOLOGY: not relevant
; MOLECULE TYPE: peptide
US-08-808-982-5

Query Match 95.0%; Score 4545.5; DB 2; Length 898;
Best Local Similarity 94.7%; Pred. No. 0;
Matches 852; Conservative 19; Mismatches 26; Indels 3; Gaps 3;
QY 1 MAVRGLWPALLGIVLAAMLWGSGAQSATVAVPGANPDLLPHFLVEPEYIVKNKP 60
|||||


```

;
; TELEFAX: (415) 343-4342
; INFORMATION FOR SEQ ID NO: 6:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 557 amino acids
; TYPE: amino acid
; STRANDEDNESS: not relevant
; TOPOLOGY: not relevant
; MOLECULE TYPE: peptide
; SEQUENCE DESCRIPTION: SEQ ID NO: 6:
US-09-306-902A-6

Query Match 58.3%; Score 2789.5; DB 3; Length 557;
Best Local Similarity 96.4%; Pred. No. 2.7e-256;
Matches 538; Conservative 2; Mismatches 15; Indels 3; Gaps 3;

QY 344 NCTSDLCVHSASGPEDVALYVGLIAVAVCLVLLLVLLVLYCRKKEGLDSVADSSILTS 403
DB 1 NCTSDLVXVHTASGPEDVALYVGLIAVAVCLVLLLVLLVLYCRKKEGLDSVADSSILTS 60

QY 404 GFQPVTSKPSKAADPHLLTITOPDLS-TTTTYQGSICPRQDGSPKFQLTNGHLLSPLGCG 462
DB 61 GFQPVTSKPSKAADPHLLTITOPDLS-TTTTYQGSICPRQDGSPKFQLTNGHLLSPLGCG 120

QY 463 RHTLHSSPTSEAEFEYSRLSTQNYFRSLPRGTNMTYGTNFNLGGRLMIPNTGISLLIP 522
DB 121 RHTLHSSPTSEAEFEYSRLSTQNYFRSLPRGTNMTYGTNFNLGGRLMIPNTGISLLIP 180

QY 523 PDAIPRGKIYIYILTLHKPEDVRLPLAGCOTLI SPIVSCGPPGVLLTRPVILAMDHCGEP 582
DB 181 PDAIPRGKIYIYILTLHKPEDVRLPLAGCOTLI SPIVSCGPPGVLLTRPVILAMDHCGEP 240

QY 583 SPDGWSLRLLKQSCBGSWEQDVHLHGEAPSHLYYQLEASACYYFTQLGRFALVGEAL 642
DB 241 SPDGWSLRLLKQSCBGSWE-DVHLHGEAPSHLYYQLEASACYYFTQLGRFALVGEAL 299

QY 643 SVAAAKRLKLLFAPVACTSLSEYNIRVYCLHDHDTDALKEVVOLEKQLGQLIQEPRVLHF 702
DB 300 SVAAAKRLKLLFAPVACTSLSEYNIRVYCLHDHDTDALKEVVOLEKQLGQLIQEPRVLHF 359

QY 703 KDSVHNLRSLSDHVPSSLWKSLLVSYQEIPIYHIWNGTQRYLHCTFTLERVSPSTSDIA 762
DB 360 XDSVHNLRSLSDHVPSSLWKSLLVSYQEIPIYHIWNGTQRYLHCTFTLERVSPSTSDIA 419

QY 763 CKLWVWQVEGQGSFISINFNITKDTFRFELLALAESEAGVPALVGSAPFKIPFLIRQKIIS 822
DB 420 CKLWVWQVEGQGSFISINFNITKDTFRFELLALAESEAGVPALVGSAPFKIPFLIRQKIIS 479

QY 823 SLDPCCRGADWRTLQAOKLHDSHLSFFASKPSPTAMILNLWEARHFPNGNLSQLAAAVA 882
DB 480 SLDPCCRGADWRTLQAOKLHDSHLSFFASKPSPTAMILNLWEARHFPNGNLSQLAAAVA 539

QY 883 GLGQPDAGLFT-VSEAE 899
DB 540 GTXPAGRWLLSQCSAE 557

RESULT 5
US-08-808-982-7
; Sequence 7, Application US/08808982
; Patent No. 5939271
; GENERAL INFORMATION:
; APPLICANT: Tessier-Lavigne, Marc
; APPLICANT: Leonardo, E. David
; APPLICANT: Hink, Lindsay
; APPLICANT: Masu, Masayuki
; APPLICANT: Kazuko, Keino-Masu
; TITLE OF INVENTION: Netrin Receptors
; NUMBER OF SEQUENCES: 8
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: SCIENCE & TECHNOLOGY LAW GROUP
; STREET: 268 BUSH STREET, SUITE 3200
; CITY: SAN FRANCISCO
; STATE: CALIFORNIA

```


Db 598 VLSPSVTCGPTGLLLCRPVVLTVPCHCAEVIAGDWIFOLKTOAHQHWB-EVVTLDDETLN 656
QY 614 HLYYCOLEASACVFTFQOLGRFALVGEALSVAAAKRLKLLLPAPVACTSLEYNIRVYCLH 673
Db 657 TPCYCOLEAKSCHILLDQGTFTVFTGESYSRAVKRLQALFAPALCTSLYSRLVYCLE 716
QY 674 DTHDALKEVVOLEKQGGOLIQEPRVLFHFKDSYHNLRLSHIDVP8SLWKSLLVSYOEIP 733
Db 717 DTPAALKEVLELERTLGGVLEPEKTLFLFKDSYHNLRLSHIDIPHAHWSKLLAKYOEIP 776
QY 734 FYHIWNGTORYLHCTFTLERVSP8SDLACKLWVWQVEGDGQSF8INFINITKDTFPAELL 793
Db 777 FYHWNGSQALHCTFTLERHSLASTFTCKVCVRQVEGEQIFOLHTTLA-ETPAGSLD 835
QY 794 ALESEAGVPAL--VGPSAFKIPFLIROKLISSIDPPCRGADWRTLAOKLHLD8HLSPPA 851
Db 836 ALCSAPGNAATTQLGPAFKIPFLIROKLISSIDPPCRGADWRTLAOKLHLD8HLSPPA 895
QY 852 SKPSTAMILNLWEARHFPNGNL8SOLA8AVAGLGP8DAGLFTV8E8EC 899
Db 896 TKASPTGVILDLWEARQDDGDLN8LASALE8MGK8EMLVAMTTD8GC 943

RESULT 6

US-09-306-902A-7
; Sequence 7, Application US/09306902A
; Patent No. 6277585
; GENERAL INFORMATION:
; APPLICANT: Tessier-Lavigne, Marc
; Leonardo, E. David
; Hink, Lindsay
; Masu, Masayuki
; Kazuko, Keino-Masu
; TITLE OF INVENTION: Netrin Receptors
; NUMBER OF SEQUENCES: 9
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: SCIENCE & TECHNOLOGY LAW GROUP
; STREET: 268 BUSH STREET, SUITE 3200
; CITY: SAN FRANCISCO
; STATE: CALIFORNIA
; COUNTRY: USA
; ZIP: 94104

COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/306,902A
; FILING DATE: 07-May-1999
; CLASSIFICATION: <Unknown>
; ATTORNEY/AGENT INFORMATION:
; NAME: OSMAN, RICHARD A
; REGISTRATION NUMBER: 36,627
; REFERENCE/DOCKET NUMBER: UC96-217
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (415) 343-4341
; TELEFAX: (415) 343-4342
; INFORMATION FOR SEQ ID NO: 7:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 943 amino acids
; TYPE: amino acid
; STRANDEDNESS: not relevant
; TOPOLOGY: not relevant
; MOLECULE TYPE: peptide
; SEQUENCE DESCRIPTION: SEQ ID NO: 7:

US-09-306-902A-7
Query Match 52.0%; Score 2490; DB 3; Length 943;
Best Local Similarity 52.3%; Pred. No. 2e-227;
Matches 496; Conservative 144; Mismatches 226; Indels 82; Gaps 19;
QY 9 PALLGIVLAALRG8GAQ8ATVANPVGANPDLLPHFLVEP8DVYI8VKN8PVL8VCKAV 68

Db 21 PSLAGI-----DSGAQ---GLPDSFP8APAEQ8PHFLLEPEDAYIV8KN8P8VELE8CRAF 70
QY 69 PATQI8FFK8G8W8V8Q8DH8V8I8R8S8TD8GS8CE8PT8ME8V8R8N8V8S8R8Q8Q8V8K8V8GL8E8Y8W8C8Q8VA 128
Db 71 PATQI8Y8K8G8W8V8S8Q8KH8VT8Q8ES8L8DE8AT8GL8R8I8R8E8V8Q8I8E8V8S8R8Q8Q8V8E8L8F8G8E8D8Y8W8C8Q8VA 130
QY 129 W8SS8GT8TK8S8K8A8Y8I8R8I8R8K8N8F8E8P8I8A8K8E8V8S8L8G8I8V8L8P8C8R8P8P8E8G8I8P8A8E8V8E8W8L8R8N8D 188
Db 131 W8SS8GT8TK8S8R8R8A8Y8I8R8I8R8K8N8F8D8E8P8I8A8K8E8V8L8L8Q8R8P8P8E8G8V8P8A8E8V8E8W8L8K8N8D 190
QY 189 LV8D8P8L8DN8V8Y8I8R8H8S8L8V8R8O8A8L8AD8T8A8N8Y8C8V8A8K8I8V8A8R8R8S8A8A8V8I8V8N8G8W8S8T8W 248
Db 191 V8I8D8P8A8Q8D8T8N8F8L8I8D8H8N8L8I8R8O8A8L8S8D8T8A8N8Y8C8V8A8K8I8V8A8R8R8S8T8A8T8I8V8V8N8G8W8S8W 250
QY 249 T8E8W8S8V8C8S8C8R8G8W8K8R8S8C8T8N8P8A8L8N8G8A8C8E8Q8N8V8H8D8R8T8V8S8L8L8V8D8G8S8W8P8S8W8K8 308
Db 251 A8E8W8S8P8C8N8R8C8R8G8W8K8R8T8C8T8N8P8A8L8N8G8A8F8C8E8Q8--A8C8Q8T8A8C8T8V8C8P8D8G8A8N8T8E8W8S8K8 309
QY 309 S8A8C8G8L8D8C8T8H8W8S8R8E8C8S8D8P8A8R8N8G8E8C8Q8T8D8L8T8R8C8T8S8D8L8C8V8--H8S8A8S8G8P8E8----- 358
Db 310 S8A8C8S8T8E8C8A8H8W8S8R8E8C8M8A8P8P8Q8N8G8R8D8C8S8T8L8D8S8X8N8C8T8D8G8L8C8V8L8N8O8R8T8L8N8D8P8K8S8R8L8E8P8S 369
QY 359 --D8V8A8Y8V8G8L8--I8A8V8C8L8V8L8L8L8V8L8V8Y8C8K8E8G8L8D8S8V8A8D8S8--L8T8S8G8F8Q8P8V8S8I8K8P8S8K8A 415
Db 370 G8D8V8A8Y8A8G8L8V8A8V8F8V8L8A8V8L8M8A8V8I8V8Y8R8N8C8R8D8F8D8T8D8S8A8A8L8T8G8F8H8F8V8N8F8K8T8A8R 429
QY 416 D8N8P8H8L8L8--T8I8Q8D8L8S8T8T8--T8Y8G8S8I8C8P8R8Q8D8P8S8K8F8Q8L8T8N8G8H8L8S8P8L8----- 459
Db 430 S8N8P8Q8L8H8P8S8A8P8D8L8T8A8S8A8G8I8Y8R8G8P8Y8A8L8Q8S8--A8D8K8I8P8M8N8S8P8L8D8P8L8S8K8I8K8Y8D8S8T8I 488
QY 460 ---G8G8-----R8H8T8H8S8P8T8S8E8A8E8F8V8S8R8L8S8T8Q8N8Y8F8R8L8P8R8G 494
Db 489 G8S8G8A8L8A8D8A8D8L8G8V8L8P8G8T8Y8P8G8F8S8R8D8T8H8L8R8S8-----A8L8G8S8Q8--H8L8G8L8P8R8D 538
QY 495 T8S8N8M8Y8T8F8N8F8L8G8R8L8M8I8P8N8T8G8I8S8I8L8I8P8D8A8I8P8R8K8I8Y8E8I8Y8L8H8K8P8D8V8L8P8L8A8--G8Q8T 553
Db 539 P8S8S8V8S8G8T8F8G8C8L8G8R8L8T8I8P8G8V8S8L8L8V8N8G8A8I8P8Q8K8F8Y8D8L8Y8L8N8K8T8E8T8--L8P8L8S8E8G8S8Q8T 597
QY 554 L8L8S8I8V8S8C8G8P8P8G8V8L8I8R8P8V8I8A8M8D8H8C8G8E8P8S8P8D8S8L8R8L8K8Q8C8E8G8S8W8E8Q8D8V8L8H8G8E8A8P8S 613
Db 598 V8L8S8P8S8V8T8C8G8P8T8G8L8L8C8R8P8V8L8T8V8P8H8C8A8E8V8I8A8D8W8I8F8O8L8K8T8O8A8H8Q8H8W8E8--E8V8T8L8D8E8T8L8N 656
QY 614 H8L8Y8C8O8L8E8A8S8A8C8Y8F8T8Q8L8G8R8F8A8L8V8G8E8A8L8S8V8A8A8K8R8L8K8L8L8F8A8P8V8A8C8T8S8L8E8Y8N8I8R8V8Y8C8L8H 673
Db 657 T8P8C8Y8C8O8L8E8A8K8S8C8H8I8L8D8Q8L8G8T8Y8V8F8T8G8E8S8Y8S8R8A8V8K8R8L8Q8A8I8F8A8P8A8L8C8T8S8L8E8Y8S8L8R8V8Y8C8L8E 716
QY 674 D8T8H8A8L8K8E8V8O8L8E8K8Q8L8G8Q8L8I8Q8E8P8R8V8L8H8F8K8D8S8Y8H8N8L8R8L8S8I8H8D8V8P8S8L8W8K8S8L8L8V8S8Y8Q8E8I8P 733
Db 717 D8T8P8A8L8K8E8V8L8E8L8E8R8T8L8G8Y8L8V8E8E8P8K8T8L8L8F8K8D8S8Y8H8N8L8R8L8S8H8D8I8P8A8H8W8R8S8K8L8A8K8Y8Q8E8I8P 776
QY 734 F8Y8H8I8W8N8G8T8O8R8Y8L8H8C8T8F8T8L8E8R8V8S8P8S8T8D8L8A8C8K8L8W8V8W8Q8V8E8G8Q8S8F8S8I8N8F8N8I8T8K8D8T8R8F8A8E8L8 793
Db 777 F8Y8H8W8N8G8S8Q8A8L8H8C8T8F8T8L8E8R8H8S8L8A8S8T8F8C8K8V8C8V8Q8V8E8G8Q8I8F8O8L8H8T8T8L8A8--E8T8P8A8S8L8D 835
QY 794 A8L8E8S8A8G8V8P8A8L8--V8G8P8A8F8K8I8P8F8I8R8O8K8I8S8S8D8P8P8C8R8R8G8A8D8W8R8T8L8A8Q8L8H8D8S8H8L8S8F8F8A 851
Db 836 A8L8C8S8A8P8G8N8A8A8T8T8Q8L8G8P8A8F8K8I8P8L8S8I8R8Q8I8C8N8S8L8D8A8P8N8S8R8G8N8D8W8L8L8A8Q8L8S8M8D8Y8L8N8Y8F8A 895
QY 852 S8K8P8S8T8A8M8I8N8L8W8E8A8R8H8F8P8N8G8N8L8S8O8L8A8A8V8A8G8L8G8O8P8A8G8L8F8T8V8S8E8A8C 899
Db 896 T8K8A8S8P8T8G8V8I8L8D8L8W8E8A8R8Q8D8D8G8D8L8N8S8L8A8S8A8L8E8M8G8K8E8M8L8V8A8M8T8D8G8C 943

RESULT 7

US-08-808-982-8
; Sequence 8, Application US/08808982
; Patent No. 5939271
; GENERAL INFORMATION:
; APPLICANT: Tessier-Lavigne, Marc
; Leonardo, E. David
; APPLICANT: Hink, Lindsay
; APPLICANT: Masu, Masayuki

APPLICANT: Kazuko, Keino-Masu
TITLE OF INVENTION: Netrin Receptors
NUMBER OF SEQUENCES: 8
CORRESPONDENCE ADDRESS:
ADDRESSEE: SCIENCE & TECHNOLOGY LAW GROUP
STREET: 268 BUSH STREET, SUITE 3200
CITY: SAN FRANCISCO
STATE: CALIFORNIA
COUNTRY: USA
ZIP: 94104
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patent In Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/808,982
FILING DATE:
CLASSIFICATION: 530
ATTORNEY/AGENT INFORMATION:
NAME: OSMAN, RICHARD A
REGISTRATION NUMBER: 36,627
REFERENCE/DOCKET NUMBER: UC96-217
TELEPHONE: (415) 343-4341
TELEFAX: (415) 343-4342
INFORMATION FOR SEQ ID NO: 8:
SEQUENCE CHARACTERISTICS:
LENGTH: 102 amino acids
TYPE: amino acid
STRANDEDNESS: not relevant
TOPOLOGY: not relevant
MOLECULE TYPE: peptide
SEQUENCE DESCRIPTION: SEQ ID NO: 8:
US-09-306-902A-8

Query Match 6.1%; Score 294; DB 2; Length 102;
Best Local Similarity 56.4%; Pred. No. 4.1e-20;
Matches 57; Conservative 16; Mismatches 28; Indels 0; Gaps 0;

QY 609 EAPSHLYYQLEASACVFTQGRFALVGEALSVAAAKRLKLLFPAPVACTSLEYNIR 668
Db 2 EETLNTPCYXQLEPRACKILLDQLGTYYFTGESYSRAVKRLQAVFAPALCTSLVSLR 61
QY 669 VYCLHDTDALKEVQLEKQGGQLIQEPRVLHFKDSYHNL 709
Db 62 VYCLEDTVALKEVLELRTLGGLVVEEPKPLMFKDSYHNL 102

RESULT 8
US-09-306-902A-8
Sequence 8, Application US/09306902A
Patent No. 6277585
GENERAL INFORMATION:
APPLICANT: Tessier-Lavigne, Marc
Leonardo, E. David
Hink, Lindsay
Masu, Masayuki
Kazuko, Keino-Masu
TITLE OF INVENTION: Netrin Receptors
NUMBER OF SEQUENCES: 9
CORRESPONDENCE ADDRESS:
ADDRESSEE: SCIENCE & TECHNOLOGY LAW GROUP
STREET: 268 BUSH STREET, SUITE 3200
CITY: SAN FRANCISCO
STATE: CALIFORNIA
COUNTRY: USA
ZIP: 94104
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patent In Release #1.0, Version #1.30
CURRENT APPLICATION DATA:

APPLICANT: Kazuko, Keino-Masu
TITLE OF INVENTION: Netrin Receptors
NUMBER OF SEQUENCES: 8
CORRESPONDENCE ADDRESS:
ADDRESSEE: SCIENCE & TECHNOLOGY LAW GROUP
STREET: 268 BUSH STREET, SUITE 3200
CITY: SAN FRANCISCO
STATE: CALIFORNIA
COUNTRY: USA
ZIP: 94104
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patent In Release #1.0, Version #1.30
CURRENT APPLICATION DATA:

APPLICANT: Kazuko, Keino-Masu
TITLE OF INVENTION: Netrin Receptors
NUMBER OF SEQUENCES: 8
CORRESPONDENCE ADDRESS:
ADDRESSEE: SCIENCE & TECHNOLOGY LAW GROUP
STREET: 268 BUSH STREET, SUITE 3200
CITY: SAN FRANCISCO
STATE: CALIFORNIA
COUNTRY: USA
ZIP: 94104
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patent In Release #1.0, Version #1.30
CURRENT APPLICATION DATA:

APPLICANT: Kazuko, Keino-Masu
TITLE OF INVENTION: Netrin Receptors
NUMBER OF SEQUENCES: 8
CORRESPONDENCE ADDRESS:
ADDRESSEE: SCIENCE & TECHNOLOGY LAW GROUP
STREET: 268 BUSH STREET, SUITE 3200
CITY: SAN FRANCISCO
STATE: CALIFORNIA
COUNTRY: USA
ZIP: 94104
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patent In Release #1.0, Version #1.30
CURRENT APPLICATION DATA:

APPLICANT: Kazuko, Keino-Masu
TITLE OF INVENTION: Netrin Receptors
NUMBER OF SEQUENCES: 8
CORRESPONDENCE ADDRESS:
ADDRESSEE: SCIENCE & TECHNOLOGY LAW GROUP
STREET: 268 BUSH STREET, SUITE 3200
CITY: SAN FRANCISCO
STATE: CALIFORNIA
COUNTRY: USA
ZIP: 94104
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patent In Release #1.0, Version #1.30
CURRENT APPLICATION DATA:

APPLICANT: Kazuko, Keino-Masu
TITLE OF INVENTION: Netrin Receptors
NUMBER OF SEQUENCES: 8
CORRESPONDENCE ADDRESS:
ADDRESSEE: SCIENCE & TECHNOLOGY LAW GROUP
STREET: 268 BUSH STREET, SUITE 3200
CITY: SAN FRANCISCO
STATE: CALIFORNIA
COUNTRY: USA
ZIP: 94104
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patent In Release #1.0, Version #1.30
CURRENT APPLICATION DATA:

APPLICANT: Kazuko, Keino-Masu
TITLE OF INVENTION: Netrin Receptors
NUMBER OF SEQUENCES: 8
CORRESPONDENCE ADDRESS:
ADDRESSEE: SCIENCE & TECHNOLOGY LAW GROUP
STREET: 268 BUSH STREET, SUITE 3200
CITY: SAN FRANCISCO
STATE: CALIFORNIA
COUNTRY: USA
ZIP: 94104
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patent In Release #1.0, Version #1.30
CURRENT APPLICATION DATA:

APPLICANT: Kazuko, Keino-Masu
TITLE OF INVENTION: Netrin Receptors
NUMBER OF SEQUENCES: 8
CORRESPONDENCE ADDRESS:
ADDRESSEE: SCIENCE & TECHNOLOGY LAW GROUP
STREET: 268 BUSH STREET, SUITE 3200
CITY: SAN FRANCISCO
STATE: CALIFORNIA
COUNTRY: USA
ZIP: 94104
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patent In Release #1.0, Version #1.30
CURRENT APPLICATION DATA:

APPLICANT: Kazuko, Keino-Masu
TITLE OF INVENTION: Netrin Receptors
NUMBER OF SEQUENCES: 8
CORRESPONDENCE ADDRESS:
ADDRESSEE: SCIENCE & TECHNOLOGY LAW GROUP
STREET: 268 BUSH STREET, SUITE 3200
CITY: SAN FRANCISCO
STATE: CALIFORNIA
COUNTRY: USA
ZIP: 94104
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patent In Release #1.0, Version #1.30
CURRENT APPLICATION DATA:

APPLICANT: Kazuko, Keino-Masu
TITLE OF INVENTION: Netrin Receptors
NUMBER OF SEQUENCES: 8
CORRESPONDENCE ADDRESS:
ADDRESSEE: SCIENCE & TECHNOLOGY LAW GROUP
STREET: 268 BUSH STREET, SUITE 3200
CITY: SAN FRANCISCO
STATE: CALIFORNIA
COUNTRY: USA
ZIP: 94104
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patent In Release #1.0, Version #1.30
CURRENT APPLICATION DATA:

APPLICANT: Kazuko, Keino-Masu
TITLE OF INVENTION: Netrin Receptors
NUMBER OF SEQUENCES: 8
CORRESPONDENCE ADDRESS:
ADDRESSEE: SCIENCE & TECHNOLOGY LAW GROUP
STREET: 268 BUSH STREET, SUITE 3200
CITY: SAN FRANCISCO
STATE: CALIFORNIA
COUNTRY: USA
ZIP: 94104
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patent In Release #1.0, Version #1.30
CURRENT APPLICATION DATA:

APPLICANT: Kazuko, Keino-Masu
TITLE OF INVENTION: Netrin Receptors
NUMBER OF SEQUENCES: 8
CORRESPONDENCE ADDRESS:
ADDRESSEE: SCIENCE & TECHNOLOGY LAW GROUP
STREET: 268 BUSH STREET, SUITE 3200
CITY: SAN FRANCISCO
STATE: CALIFORNIA
COUNTRY: USA
ZIP: 94104
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patent In Release #1.0, Version #1.30
CURRENT APPLICATION DATA:

APPLICANT: Kazuko, Keino-Masu
TITLE OF INVENTION: Netrin Receptors
NUMBER OF SEQUENCES: 8
CORRESPONDENCE ADDRESS:
ADDRESSEE: SCIENCE & TECHNOLOGY LAW GROUP
STREET: 268 BUSH STREET, SUITE 3200
CITY: SAN FRANCISCO
STATE: CALIFORNIA
COUNTRY: USA
ZIP: 94104
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patent In Release #1.0, Version #1.30
CURRENT APPLICATION DATA:

APPLICANT: Kazuko, Keino-Masu
TITLE OF INVENTION: Netrin Receptors
NUMBER OF SEQUENCES: 8
CORRESPONDENCE ADDRESS:
ADDRESSEE: SCIENCE & TECHNOLOGY LAW GROUP
STREET: 268 BUSH STREET, SUITE 3200
CITY: SAN FRANCISCO
STATE: CALIFORNIA
COUNTRY: USA
ZIP: 94104
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patent In Release #1.0, Version #1.30
CURRENT APPLICATION DATA:

APPLICANT: Kazuko, Keino-Masu
TITLE OF INVENTION: Netrin Receptors
NUMBER OF SEQUENCES: 8
CORRESPONDENCE ADDRESS:
ADDRESSEE: SCIENCE & TECHNOLOGY LAW GROUP
STREET: 268 BUSH STREET, SUITE 3200
CITY: SAN FRANCISCO
STATE: CALIFORNIA
COUNTRY: USA
ZIP: 94104
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patent In Release #1.0, Version #1.30
CURRENT APPLICATION DATA:

APPLICANT: Kazuko, Keino-Masu
TITLE OF INVENTION: Netrin Receptors
NUMBER OF SEQUENCES: 8
CORRESPONDENCE ADDRESS:
ADDRESSEE: SCIENCE & TECHNOLOGY LAW GROUP
STREET: 268 BUSH STREET, SUITE 3200
CITY: SAN FRANCISCO
STATE: CALIFORNIA
COUNTRY: USA
ZIP: 94104
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patent In Release #1.0, Version #1.30
CURRENT APPLICATION DATA:

APPLICANT: Kazuko, Keino-Masu
TITLE OF INVENTION: Netrin Receptors
NUMBER OF SEQUENCES: 8
CORRESPONDENCE ADDRESS:
ADDRESSEE: SCIENCE & TECHNOLOGY LAW GROUP
STREET: 268 BUSH STREET, SUITE 3200
CITY: SAN FRANCISCO
STATE: CALIFORNIA
COUNTRY: USA
ZIP: 94104
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patent In Release #1.0, Version #1.30
CURRENT APPLICATION DATA:

APPLICANT: Kazuko, Keino-Masu
TITLE OF INVENTION: Netrin Receptors
NUMBER OF SEQUENCES: 8
CORRESPONDENCE ADDRESS:
ADDRESSEE: SCIENCE & TECHNOLOGY LAW GROUP
STREET: 268 BUSH STREET, SUITE 3200
CITY: SAN FRANCISCO
STATE: CALIFORNIA
COUNTRY: USA
ZIP: 94104
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patent In Release #1.0, Version #1.30
CURRENT APPLICATION DATA:

APPLICANT: Kazuko, Keino-Masu
TITLE OF INVENTION: Netrin Receptors
NUMBER OF SEQUENCES: 8
CORRESPONDENCE ADDRESS:
ADDRESSEE: SCIENCE & TECHNOLOGY LAW GROUP
STREET: 268 BUSH STREET, SUITE 3200
CITY: SAN FRANCISCO
STATE: CALIFORNIA
COUNTRY: USA
ZIP: 94104
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patent In Release #1.0, Version #1.30
CURRENT APPLICATION DATA:

APPLICANT: Kazuko, Keino-Masu
TITLE OF INVENTION: Netrin Receptors
NUMBER OF SEQUENCES: 8

US-08-313-288B-19

Query Match	5.7%	Score 273.5;	DB 1;	Length 1172;
Best Local Similarity	28.9%;	Pred. No. 2.le-16;		
Matches	74; Conservative	27; Mismatches 110;	Indels 45; Gaps	8;
QY	209	RQARLADTANYTCVAKNIVARRRSASAA-VIVYVNGGWSTWTEWSVCSCGRCGWQKRSR	267	
Db	403	: QGRSCDVTSNTCLGPSIQTRACLSLKCDRIQRDGGCWSHPSSCSVTGCGVNITIR	462	
QY	268	SCINPAPLNGGARCEQNVDHRTVSSLVSDGSWSPWKSASCGLDCT--HWRSRECS	324	
Db	463	LCMSPVPQMGGKNCKSGSRETCAQQCAPCIDGRWSFSPSWSACTVTCAGGIERTVCN	522	
QY	325	DPAPRNGBECQGTDLJDTNRCTSDLCVHSASGEDVALYYGLIAVAACLVLILLLVILTV	384	
Db	523	SPEFYGGKACVGDVQERQMCNKRS-----PVDGCLSNPCFPGAQC-----	564	
QY	385	CRKKEGLDSADVASSIILTSGFQPUSI--KPSKADNPHELLIIQDDLSTTT-----TYQ	434	
Db	565	-----SSFPGDS-WSCGFPCPVFLGNTHCEDLDDECALVPDICSTSKVPCVNTQP	615	
QY	435	GSLC-----PRQDGSP	446	
Db	616	GFHCLPCPPRYRGNOP	631	

```

RESULT 10
US-08-985-526-3
; Sequence 3, Application US/08985526
; Patent No. 6080728
; GENERAL INFORMATION:
; APPLICANT: Mixson, James A
; TITLE OF INVENTION: CARRIER:DNA COMPLEXES CONTAINING DNA
; TITLE OF INVENTION: ENCODING ANTI-ANGIOGENIC PEPTIDES AND THEIR USE IN GENE
; TITLE OF INVENTION: THERAPY
; NUMBER OF SEQUENCES: 43
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Connolly, Bove, Lodge, & Hutz
; STREET: 1220 Market Street, P.O. Box 2207
; CITY: Wilmington
; STATE: Delaware
; COUNTRY: U.S.A.
; ZIP: 19899
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent In Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/985,526
; FILING DATE:
; CLASSIFICATION:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/608,845
; FILING DATE: 16-JUL-1996
; ATTORNEY/AGENT INFORMATION:
; NAME: McMorow Jr., Robert G
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (302) 658-9141
; TELEFAX: (302) 658-5613
; INFORMATION FOR SEQ ID NO: 3:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 441 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; US-08-985-526-3

```

```

132 FKQDGGSHWSPWSSCSVTCDGQVITRITLQNSPSPQMGKPCGEARETKACKDACP 191
114 EKVGLEEYWCQAWSSSGTTKSKAVIRIARLKNFEQBLAKEVSLQGVILPCRPP 173
192 NGMGPSWPDICSVTCGGGVQKRSRLCV---DSMTENKELANELR-----RPP 239
174 ---EGIPPAEVEUARNED-LVDPSLDPNVYITREHSLVVRQARLADTANYTCVAKNIVA 228
240 LCYHNG-----VOYRNEEWIVDSCTE-----CHQNSVT 269
229 RRRSASAAVIVYNG-----GWSTWTWVSVCSACGRGWQKESRSC 269
270 ICKVSCFIMPSCSNATVPDGECCPRCWPDSADGWSPWSEWTSCTSCGNGIQQRGRSC 329
270 TNPAPLGGAFCEGONVHDTV-----SSLLYSVQGSWSPWKSWSACGLDC---THWRSRE 322
330 DS-----LNNR---CEGSSVOTRCHIQECDKRFQDGGWSHWSPWSSCSVTCDGQVITRITL 384
323 CSDPAPRNGBECQGTDLDRNCTSDLC-VHSASGP 357
385 CNSPSPQMGKPCGEARETKACKDACPINGWGP 420

RESULT 11
PCT-US93-01652-1
; Sequence 1, Application PC/TUS9301652
; GENERAL INFORMATION:
; APPLICANT: Bouck, Noel P.
; APPLICANT: Polverini, Peter J.
; APPLICANT: Good, Deborah J.
; APPLICANT: Frazier, William A.
; TITLE OF INVENTION: Method and Composition for
; TITLE OF INVENTION: Inhibiting Angiogenesis
; NUMBER OF SEQUENCES: 12
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Iilton, Fallon, Lungmus & Chestnut
; STREET: 100 South Wacker Drive, Suite 960
; CITY: Chicago
; STATE: Illinois
; COUNTRY: USA
; ZIP: 60606-4002
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: PCT/US93/01652
; FILING DATE: 19930222
; CLASSIFICATION:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US/07/841,656
; FILING DATE: 24-FEB-1992
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US/07/464,369
; FILING DATE: 12-JAN-1990
; ATTORNEY/AGENT INFORMATION:
; NAME: Fentress, Susan B.
; REGISTRATION NUMBER: 31,327
; REFERENCE/DOCKET NUMBER: 92005-PCT
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (312)-456-8000
; TELEFAX: (312)-456-7776
; INFORMATION FOR SEQ ID NO: 1:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 239 amino acids
; TYPE: AMINO ACID
; STRANDEDNESS: unknown
; TOPOLOGY: unknown
; MOLECULE TYPE: peptide
PCT-US93-01652-1

```


; CURRENT FILING DATE: 2000-03-31
; PRIOR APPLICATION NUMBER: 60/065,544
; PRIOR FILING DATE: 1997-11-14
; PRIOR APPLICATION NUMBER: 60/081,057
; PRIOR FILING DATE: 1998-04-07
; NUMBER OF SEQ ID NOS: 20
; SOFTWARE: Patent In Ver. 2.0
; SEQ ID NO 15
; LENGTH: 1395
; TYPE: PRT
; ORGANISM: Drosophila melanogaster
US-09-540-245A-15

Query Match 4.9%; Score 234; DB 3; Length 1395;
Best Local Similarity 20.5%; Pred. No. 1.6e-12;
Matches 188; Conservative 103; Mismatches 262; Indels 364; Gaps 46;

QY 4 RPLGPAALIGIVLAALRGGAQASATVANPFGA-----NPDLLPHFLVPEDEYVIVKN 58
DB 28 RMWLLPAWLLLVLA-----SNGLPVAVRGVQSPRIIEH---PTDLVVKKN 70

QY 59 KPVLIVCK--AVPATQIFKNGEVRQVDHVIERSDGSSEPTMEVRINVS- 110
DB 71 EPATLNCKVEGPEPTI-----EWFK-----DGEPT---VSTNEKKSHRVQF 108

QY 111 -----QVEKVFGLBEYKCOQVAVSSSGTTKSKAYIRIARLRKNFEQELAKEV 160
DB 109 KDGALFFRTMGKEQDGEY--CVAKNRVQAVSHASQIAVLRRDPRVPRDTRV 166

QY 161 SLEQGIPLPCRPPEGIPPAEVEWLRN-----EDLVDPSL--DENVTITREHSLVVRQARLA 214
DB 167 AKGETALLECQPKGIPETLWIKDGVPLDLKAMSGASSRVRIVDGNLLISNVEPI 226

QY 215 DTANTYCVAKNIVARRRSASAIVIVVN-----GWSSTWTEWSVCASGR 260
DB 227 DEGNKYKIAQNLVGTRESSYAKLIQVKEPYFMKEPRKDQVMLYQOTATF-----HCSV 278

QY 261 G-----WQK-----RSRCTNPAPLNGGAF-CEGON----- 285
DB 279 GGDPPPKVLWKEBGNIPVSRARILHDEKSLISNITPDEGYVCEAHNNVGQISARAS 338

QY 286 --VH-----DRTVSSILLVSDG-----WSWP-----WSKWSACGL--DCTHWRSE 322
DB 339 LIVHAPPNETKPSNKKVGLNGVQLPCMASGNPPSPVFTKEGVTLMFPNSSHGRQV 398

QY 323 CSDPAPRNGEECQGTDLTRNCTSDLCVHSAGSEPDVALYVGLIAVAVCLVLLVLL 382
DB 399 AADGT-----LQITDV-----RQDEGY--VCSAF----- 422

QY 383 VYCRKKEGLSDVADSSILTSFGFQVSIKPSKADN--PHLLTIQPDLSITTYQGLS--- 437
DB 423 -----SVDSSTVR-----VFLQVSVDERPPPIIQGP--ANQTLPGKSVATL 464

QY 438 -CFQDGPSPKFQ-LTNGHLLPLGGRHTLHSSPTSSEAEFVSRLSNTQYFRSLPRGT 495
DB 465 PCRAATGNPSPRIKWFHDGHAQ--AGNRYSIIQSSLRVDDQLQSDSGYTTCTASGERGE 522

QY 496 SNTYGTENFLGRLMPTNGISLLIPDPAIPRGKIYEIYLHLKPEDVRLPLAGCQTL 555
DB 523 TS-----WQK-----WAATLVEKPGSTSLHRA----- 543

QY 556 SPIVSCGPPGVLLTRPVILAMDHCGEPPSPDSLSLR-LKKQSCGS----- 599
DB 544 DPSTVPAPPGT-----PKVLNV-----SRTSISLRWAKSQEKPAGVGIITYVEYFSPD 593

QY 600 -----WEQDVLHLGEEAPSHLYYQLEASACYVF-----TEOLGR 634
DB 594 LOTGMIVAAHRVGD---TQVTTISGLTPGTSYVFLVRAENTQGISVPSGLSNVIKTIADF 650

QY 635 FALVGEALSVA---AAKRLKLLFPAPVACTSLEYNIRVYCLHDTHALKVYVQLEKQLG 690
DB 651 DAASANDLSAARTLLTGKVELIDASAINAVRLIE---WMLHVSAD-----EKYVE 699

QY 691 GQLIQEPRVLHFKOS-----YHNLR-----SIHVD 716
DB 700 GLRI-----HYKASVESAQYHSITVMDASAESFVGNLKKYKYEPFLTPFFETIEQ 753

QY 717 PSSLWKSLLVSQEP 733
DB 754 PSN---SKTALTYEDVP 767

RESULT 15
US-09-540-245A-18
; Sequence 18, Application US/09540245A
; Patent No. 6270984
; GENERAL INFORMATION:
; APPLICANT: Goodman, Corey
; APPLICANT: Kid, Thomas
; APPLICANT: Brose, Katja
; APPLICANT: Tessier-Lavigne, Marc
; TITLE OF INVENTION: Modulating Robo: Ligand Interactions
; FILE REFERENCE: B98-031-3
; CURRENT APPLICATION NUMBER: US/09/540,245A
; CURRENT FILING DATE: 2000-03-31
; PRIOR APPLICATION NUMBER: 60/065,544
; PRIOR FILING DATE: 1997-11-14
; PRIOR APPLICATION NUMBER: 60/081,057
; PRIOR FILING DATE: 1998-04-07
; NUMBER OF SEQ ID NOS: 20
; SOFTWARE: Patent In Ver. 2.0
; SEQ ID NO 18
; LENGTH: 1651
; TYPE: PRT
; ORGANISM: human
US-09-540-245A-18

Query Match 4.5%; Score 216.5; DB 3; Length 1651;
Best Local Similarity 20.7%; Pred. No. 1e-10;
Matches 213; Conservative 139; Mismatches 381; Indels 295; Gaps 54;

QY 41 DLLPHFLVEPEDEYVIVKNPVLIVCKAV--PATQIFKNGEVRQVDHVIERSDGS 98
DB 65 DFPFRIVEHPSDLIVSKGEPAFLNCKAEGRPPTI-----EWYKGERVETKDDPRSH 118

QY 99 EPTME-----VRINVSQVEKVFGLBEYKCOQVAVSSSGTTKSKAYIRIARLRKNFE 152
DB 119 RMLPSSGSLFFLIRVHGRKSRP-----DEGVTVCVARNYLGEAVSHNASLEVALRDDPR 173

QY 153 QBLAKESLEQGIPLPCRPPEGIPPAEVEWLRNEDLVDPSLDPNVYIT-REHSLVVRQA 211
DB 174 QNPSDVMVAVGPAVMEQPPRGHPEPTISWKKGSPLD---DKDERITIRGGKLMITYT 230

QY 212 RLADTANTYCVAKNIVARRRSASAIVIVVNGW-----STWTEWSV---CSASCGRG 261
DB 231 RKSDAGKVCVGTNVMGERESEVABLTVLERFSVKRPSNLAVTVDDSAEFKCEA---RG 287

QY 262 -----WQK-----RSRCTNPAPLNGGAFCEGQVNHDRITVSSLLV 296
DB 288 DPVPTVRKDGELPKSEYREIRDDHTLKIRKVT--AGDMGSYTCVAENMVKAASATL 345

QY 297 SYDGS-----WSPWSKWSACGLDCTHWRSECDPAP-----RNGGEEC-----QGTDL 340
DB 346 TVQEPFHFVVKPRDQVVALGRVT--FQCEATGNPQPAIFWRREGSQNLLFSYQPQSSSR 404

QY 341 DTRNCTSDLCVHSAGSEPDVALYVGLIAVAVCLVLLVLLVLYCRKKEGLSDVADSSI 400
DB 405 FSVSQTGLTITNVQ-RSDVGYI-----ICQTL-----NVA-GSI 437

QY 401 LTSFGQPVSIKSKADNPHLLTIQPDLSITTYQGLS-----CFRQDGPSPKF----- 448
DB 438 ITKAY--LEVTVIADRPVPPVIRQGPVQVAVDGTQFVLSVATGSPVETILWRKDVLY 495

QY 449 -----QLTNGHL-----LSPLGG-GRHTLHSSPTSSEAEFVSRLSNTQYFRSLPRGTS 496
DB 496 STQDSRIKOLENGVLQIRYAKLGDITGRYTCIASTPSGEA----- 534

Search completed: October 4, 2004, 18:33:20
Job time : 38 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: October 4, 2004, 18:29:07 ; Search time 66 Seconds
(without alignments)
4383.295 Million cell updates/sec

Title: US-09-970-944-2

Perfect score: 4787

Sequence: 1 MAVRGLWPALLGIVLAWL.....AVAGLGQDPAGLFTVSEABC 899

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 1351062 seqs, 32179191 residues

Total number of hits satisfying chosen parameters: 1351062

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : Published Applications AA:

- 1: /cgn2_6/ptodata/2/pubpaa/US07_PUBCOMB.pep.*
- 2: /cgn2_6/ptodata/2/pubpaa/PCT_NEW_PUB.pep.*
- 3: /cgn2_6/ptodata/2/pubpaa/US06_NEW_PUB.pep.*
- 4: /cgn2_6/ptodata/2/pubpaa/US06_PUBCOMB.pep.*
- 5: /cgn2_6/ptodata/2/pubpaa/US07_NEW_PUB.pep.*
- 6: /cgn2_6/ptodata/2/pubpaa/PCTUS_PUBCOMB.pep.*
- 7: /cgn2_6/ptodata/2/pubpaa/US08_NEW_PUB.pep.*
- 8: /cgn2_6/ptodata/2/pubpaa/US08_PUBCOMB.pep.*
- 9: /cgn2_6/ptodata/2/pubpaa/US09A_PUBCOMB.pep.*
- 10: /cgn2_6/ptodata/2/pubpaa/US09B_PUBCOMB.pep.*
- 11: /cgn2_6/ptodata/2/pubpaa/US09C_PUBCOMB.pep.*
- 12: /cgn2_6/ptodata/2/pubpaa/US09_NEW_PUB.pep.*
- 13: /cgn2_6/ptodata/2/pubpaa/US10A_PUBCOMB.pep.*
- 14: /cgn2_6/ptodata/2/pubpaa/US10B_PUBCOMB.pep.*
- 15: /cgn2_6/ptodata/2/pubpaa/US10C_PUBCOMB.pep.*
- 16: /cgn2_6/ptodata/2/pubpaa/US10_NEW_PUB.pep.*
- 17: /cgn2_6/ptodata/2/pubpaa/US60_NEW_PUB.pep.*
- 18: /cgn2_6/ptodata/2/pubpaa/US60_PUBCOMB.pep.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	4787	100.0	899	11	US-09-970-944-2
2	4698.5	98.2	898	12	US-09-918-779-2
3	4698.5	98.2	898	16	US-10-624-932-2
4	4545.5	95.0	898	10	US-09-933-261-5
5	4545.5	95.0	898	11	US-09-970-944-13
6	4545.5	95.0	898	14	US-10-256-702-5
7	4545.5	95.0	898	14	US-10-240-154-16
8	4379.5	91.5	842	16	US-10-311-623-1
9	2819	58.9	544	11	US-09-970-944-14
10	2789.5	58.3	557	10	US-09-933-261-6
11	2789.5	58.3	557	14	US-10-256-702-6
12	2705.5	56.5	931	11	US-09-970-944-15
13	2705.5	56.5	931	11	US-09-970-944-35
14	2705.5	56.5	931	12	US-09-972-211-121
15	2705.5	56.5	931	12	US-10-037-417-117

16	2705.5	56.5	931	12	US-10-096-625-121
17	2705.5	56.5	1010	12	US-10-218-779-35
18	2680.5	56.0	931	11	US-09-970-944-16
19	2680.5	56.0	931	12	US-09-972-211-125
20	2680.5	56.0	931	12	US-10-096-625-125
21	2677.5	55.9	931	12	US-10-037-417-120
22	2673.5	55.8	931	11	US-09-970-944-17
23	2673.5	55.8	931	12	US-10-087-684-36
24	2673.5	55.8	931	12	US-10-218-779-36
25	2673.5	55.8	931	12	US-09-972-211-122
26	2673.5	55.8	931	12	US-10-037-417-118
27	2673.5	55.8	931	12	US-10-037-417-119
28	2673.5	55.8	931	12	US-10-096-625-122
29	2497	52.2	945	12	US-10-087-684-34
30	2497	52.2	945	12	US-10-218-779-34
31	2497	52.2	945	12	US-09-972-211-124
32	2497	52.2	945	12	US-10-037-417-121
33	2497	52.2	945	12	US-10-096-625-124
34	2493	52.1	945	12	US-10-087-684-33
35	2493	52.1	945	12	US-10-218-779-33
36	2493	52.1	945	12	US-09-972-211-123
37	2493	52.1	945	12	US-10-096-625-123
38	2490.5	52.0	946	12	US-09-972-211-56
39	2490.5	52.0	946	12	US-10-096-625-56
40	2490	52.0	943	10	US-09-933-261-7
41	2490	52.0	943	14	US-10-256-702-7
42	2484	51.9	933	12	US-10-087-684-2
43	2484	51.9	933	12	US-10-087-684-4
44	2484	51.9	933	12	US-10-218-779-2
45	2484	51.9	933	12	US-10-218-779-4

ALIGNMENTS

RESULT 1

US-09-970-944-2
; Sequence 2, Application US/09970944
; Publication No. US20030204052A1
; GENERAL INFORMATION:
; APPLICANT: Herrman, John L
; APPLICANT: Shimkets, Richard A
; TITLE OF INVENTION: No. US20030204052A1el Proteins and Nucleic Acids Encoding Same an
; TITLE OF INVENTION: Antibodies Directed Against these Proteins
; FILE REFERENCE: 21402-138
; CURRENT APPLICATION NUMBER: US/09/970,944
; PRIOR FILING DATE: 2002-05-02
; PRIOR FILING DATE: 2000-10-04
; NUMBER OF SEQ ID NOS: 62
; SOFTWARE: Patent In Ver. 2.1
; SEQ ID NO 2
; LENGTH: 899
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-970-944-2

Query Match	100.0%	Score 4787;	DB 11;	Length 899;
Best Local Similarity	100.0%;	Pred. No. 0;		
Matches 899;	Conservative 0;	Mismatches 0;	Indels 0;	Gaps 0;
Qy	1	MAVRPGLWPALLGIVLAWL	RGSGAQGSATVANVPGANPDLLPHFLVPEPDVYIVKNKP	60
Db	1	MAVRPGLWPALLGIVLAWL	RGSGAQGSATVANVPGANPDLLPHFLVPEPDVYIVKNKP	60
Qy	61	VLLVCKAVPATQIFKNGEWMVRQVDHVIERTDSSGSEPTMEYRINVSROOVKEVFGLE	120	
Db	61	VLLVCKAVPATQIFKNGEWMVRQVDHVIERTDSSGSEPTMEYRINVSROOVKEVFGLE	120	
Qy	121	EYWCQCVAWSSGGTTKQKAYIRIARLRKNFQEQPLAKEVSLQGIPLPCPPGIPPAE	180	
Db	121	EYWCQCVAWSSGGTTKQKAYIRIARLRKNFQEQPLAKEVSLQGIPLPCPPGIPPAE	180	

QY 181 VEWLNEDLVDPSPDNVYITREHSLVVRQARLADTANYTCVAKNIVARRRSASAIVY 240
Db 181 VEWLNEDLVDPSPDNVYITREHSLVVRQARLADTANYTCVAKNIVARRRSASAIVY 240
QY 241 VNGGWSTWTSVCSASCGRGWQKRSCTNPAPLNGGAFCEGQNVHRTVSSLLVSDG 300
Db 241 VNGGWSTWTSVCSASCGRGWQKRSCTNPAPLNGGAFCEGQNVHRTVSSLLVSDG 300
QY 301 SWSPMSKWSACGLDCTHWRSECSDPAPNGGECQGTDLTRNCTSDLCVHSASGPDV 360
Db 301 SWSPMSKWSACGLDCTHWRSECSDPAPNGGECQGTDLTRNCTSDLCVHSASGPDV 360
QY 361 ALYVGLIAVAVCLVLLVLLVILVYCRKKEGLSDVADSSILTSFGQPSIKPSKADNPHL 420
Db 361 ALYVGLIAVAVCLVLLVLLVILVYCRKKEGLSDVADSSILTSFGQPSIKPSKADNPHL 420
QY 421 LTIQPDLSLTYYYQSLCPRODGPSPKFQLTNGHLLSPGGGRHTLHSSPTSSEAEFVS 480
Db 421 LTIQPDLSLTYYYQSLCPRODGPSPKFQLTNGHLLSPGGGRHTLHSSPTSSEAEFVS 480
QY 481 RLSTONYFRSLPRGTSNNYTGTFNPLGRLMIPNTGISILLIPDDAI PRGKIYEIYLT LHK 540
Db 481 RLSTONYFRSLPRGTSNNYTGTFNPLGRLMIPNTGISILLIPDDAI PRGKIYEIYLT LHK 540
QY 541 PEDVRLPLAGCOTLLSPVSCGPGVLTTRPVILAMDHCGEPPDSWSILRLKKQCEGSM 600
Db 541 PEDVRLPLAGCOTLLSPVSCGPGVLTTRPVILAMDHCGEPPDSWSILRLKKQCEGSM 600
QY 601 EQDVLHGEAPSHLYYQLEASACVYFTEQGRPALVGEALSVAAXRLKLLFPAPVAC 660
Db 601 EQDVLHGEAPSHLYYQLEASACVYFTEQGRPALVGEALSVAAXRLKLLFPAPVAC 660
QY 661 TSLEYNIRVYCLHDTLHDKALKEVQLEKQGLQLOEPRVLPKDSYHNLRLSIHDPSSL 720
Db 661 TSLEYNIRVYCLHDTLHDKALKEVQLEKQGLQLOEPRVLPKDSYHNLRLSIHDPSSL 720
QY 721 WSKLLVSYQETPFVHIWNGTORYLHCTFTLERSVPSTSDLACKLWVQVEGDGOSFSIN 780
Db 721 WSKLLVSYQETPFVHIWNGTORYLHCTFTLERSVPSTSDLACKLWVQVEGDGOSFSIN 780
QY 781 FNITKDTFABELLASEAGVPALVGPFAKIPFLIRQKIISLPPCRRGADMTLQAK 840
Db 781 FNITKDTFABELLASEAGVPALVGPFAKIPFLIRQKIISLPPCRRGADMTLQAK 840
QY 841 LHLDSLSPFASKSPPTAMILNWEARHPNGNLSQLAAVAGLQPDAGLFTVSEAE 899
Db 841 LHLDSLSPFASKSPPTAMILNWEARHPNGNLSQLAAVAGLQPDAGLFTVSEAE 899

RESULT 2

US-09-918-779-2
; Sequence 2, Application US/09918779
; Publication No. US2003064369A1
; GENERAL INFORMATION:
; APPLICANT: Taupier, Raymond
; APPLICANT: Padigar, Muralidhara
; APPLICANT: Rastelli, Luca
; APPLICANT: Spaderna, Steven
; APPLICANT: Shimkets, Richard
; APPLICANT: Zerkhusen, Bryan
; APPLICANT: Spytek, Kimberly
; APPLICANT: Shenoy, Suresh
; APPLICANT: Li, Li
; APPLICANT: Gusev, Vladimir
; APPLICANT: Grosse, William
; APPLICANT: Alsobrook, John
; APPLICANT: Lepley, Denise
; APPLICANT: Burgess, Catherine
; APPLICANT: Gerlach, Valerie
; APPLICANT: Ellerman, Karen
; APPLICANT: MacDougall, John
; APPLICANT: Stone, David

; APPLICANT: Smithson, Glenda
; TITLE OF INVENTION: Novel Proteins and Nucleic Acids Encoding Same
; FILE REFERENCE: 21402-074 US
; CURRENT APPLICATION NUMBER: US/09/918,779
; CURRENT FILING DATE: 2001-07-30
; PRIOR APPLICATION NUMBER: 60/221,409
; PRIOR FILING DATE: 2000-07-28
; PRIOR APPLICATION NUMBER: 60/222,840
; PRIOR FILING DATE: 2000-08-04
; PRIOR APPLICATION NUMBER: 60/223,752
; PRIOR FILING DATE: 2000-08-08
; PRIOR APPLICATION NUMBER: 60/223,762
; PRIOR FILING DATE: 2000-08-08
; PRIOR APPLICATION NUMBER: 60/223,770
; PRIOR FILING DATE: 2000-08-08
; PRIOR APPLICATION NUMBER: 60/223,769
; PRIOR FILING DATE: 2000-08-08
; PRIOR APPLICATION NUMBER: 60/225,146
; PRIOR FILING DATE: 2000-08-14
; PRIOR APPLICATION NUMBER: 60/225,392
; PRIOR FILING DATE: 2000-08-15
; PRIOR APPLICATION NUMBER: 60/225,470
; PRIOR FILING DATE: 2000-08-15
; PRIOR APPLICATION NUMBER: 60/225,697
; PRIOR FILING DATE: 2000-08-16
; PRIOR APPLICATION NUMBER: 60/263,662
; PRIOR FILING DATE: 2001-02-01
; PRIOR APPLICATION NUMBER: 60/281,645
; PRIOR FILING DATE: 2001-04-05
; NUMBER OF SEQ ID NOS: 61
; SOFTWARE: Patent In Ver. 2.1
; SEQ ID NO 2
; LENGTH: 898
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-918-779-2

Query Match 98.2%; Score 4698.5; DB 12; Length 898;
Best Local Similarity 98.7%; Pred. No. 0;
Matches 888; Conservative 2; Mismatches 7; Indels 3; Gaps 3;
QY 1 MAVRGLWPALLGIVLAALRGSGAQQSATVANPVPANPDLLPHFLVEPEDVYVKNKP 60
Db 1 MAVRGLWPALLGIVLAALRGSGAQQSATVANPVPANPDLLPHFLVEPEDVYVKNKP 60
QY 61 VLLVCKAVPATQIPKCNGEWVRQVDHVIERSDTGSSGSEPTMEVRINVSROQVEKVFGL 120
Db 61 VLLVCKAVPATQIPKCNGEWVRQVDHVIERSDTGSSGSEPTMEVRINVSROQVEKVFGL 120
QY 121 EYWCOCVAVSSGTTKSKAVIRIARLKNFEQEPLEKEVSLQEGIVLPCRPPEGIPPAE 180
Db 121 EYWCOCVAVSSGTTKSKAVIRIARLKNFEQEPLEKEVSLQEGIVLPCRPPEGIPPAE 180
QY 181 VEWLNEDLVDPSPDNVYITREHSLVVRQARLADTANYTCVAKNIVARRRSASAIVY 240
Db 181 VEWLNEDLVDPSPDNVYITREHSLVVRQARLADTANYTCVAKNIVARRRSASAIVY 240
QY 241 VNGGWSTWTSVCSASCGRGWQKRSCTNPAPLNGGAFCEGQNVHRTVSSLLVSDG 300
Db 241 VNGGWSTWTSVCSASCGRGWQKRSCTNPAPLNGGAFCEGQNVHRTVSSLLVSDG 300
QY 301 SWSPMSKWSACGLDCTHWRSECSDPAPNGGECQGTDLTRNCTSDLCVHSASGPDV 360
Db 301 SWSPMSKWSACGLDCTHWRSECSDPAPNGGECQGTDLTRNCTSDLCVHSASGPDV 360
QY 361 ALYVGLIAVAVCLVLLVLLVILVYCRKKEGLSDVADSSILTSFGQPSIKPSKADNPHL 420
Db 361 ALYVGLIAVAVCLVLLVLLVILVYCRKKEGLSDVADSSILTSFGQPSIKPSKADNPHL 420
QY 421 LTIQPDLSLTYYYQSLCPRODGPSPKFQLTNGHLLSPGGGRHTLHSSPTSSEAEFVS 479
Db 421 LTIQPDLSLTYYYQSLCPRODGPSPKFQLTNGHLLSPGGGRHTLHSSPTSSEAEFVS 479

```
QY 480 SRLSTQNYFRSLPRGTSNNMTYGTFFNLGGLMLPNTGISLILPPDAIPRGKIYIYLTLLH 539
Db 480 SRLSTQNYFRSLPRGTSNNMTYGTFFNLGGLMLPNTGISLILPPDAIPRGKIYIYLTLLH 539
QY 540 KPEDVRLPLAGCQTLSPVSCGPPGVLLTRPVILAMDHGCEPSDPSWSLRLKKQCEGS 599
Db 540 KPEDVRLPLAGCQTLSPVSCGPPGVLLTRPVILAMDHGCEPSDPSWSLRLKKQCEGS 599
QY 600 WEQDVLHLGEEAPSHLYYCOLEASACVFTFEQGRFALVGEALSVAAAKLKLFPAPVA 659
Db 600 WE-DVLHLGEEAPSHLYYCOLEASACVFTFEQGRFALVGEALSVAAAKLKLFPAPVA 658
QY 660 CTSLEYNIRVYCLDTHDALKEVVQLEKQGGOLIOBPRVLHFKDSVHNLRLSIHDPSS 719
Db 659 CTSLEYNIRVYCLDTHDALKEVVQLEKQGGOLIOBPRVLHFKDSVHNLRLSIHDPSS 718
QY 720 LKSKLLVSYQEIIPFFYHINWGTQRYLHCTFTLBRVSPSTSDACKLWVWQVEGDGQSF 779
Db 719 LKSKLLVSYQEIIPFFYHINWGTQRYLHCTFTLBRVSPSTSDACKLWVWQVEGDGQSF 778
QY 780 NFNITKDTFAELLALAESEAGVPALVGPSPAFKIPFLIRQKIISLDPCCRRGADWRTLAQ 839
Db 779 NFNITKDTFAELLALAESEAGVPALVGPSPAFKIPFLIRQKIISLDPCCRRGADWRTLAQ 838
QY 840 KLHLDHSLSPFASKPSPMTAMLNWEARHPNGNLSQLAAGVAGLGQPDAGLFTVSEAC 899
Db 839 KLHLDHSLSPFASKPSPMTAMLNWEARHPNGNLSQLAAGVAGLGQPDAGLFTVSEAC 898

RESULT 3
US-10-624-932-2
; Sequence 2, Application US/10624932
; Publication No. US20040096877A1
; GENERAL INFORMATION:
; APPLICANT: Taupier, Raymond
; APPLICANT: Padigaru, Muralidhara
; APPLICANT: Rastelli, Luca
; APPLICANT: Spaderna, Steven
; APPLICANT: Shimkets, Richard
; APPLICANT: Zerhusen, Bryan
; APPLICANT: Spytek, Kimberly
; APPLICANT: Shenoy, Suresh
; APPLICANT: Li, Li
; APPLICANT: Gusev, Vladimir
; APPLICANT: Grosse, William
; APPLICANT: Alsobrook, John
; APPLICANT: Lepley, Denise
; APPLICANT: Burgess, Catherine
; APPLICANT: Gerlach, Valerie
; APPLICANT: Ellerman, Karen
; APPLICANT: MacDougall, John
; APPLICANT: Stone, David
; APPLICANT: Smithson, Glenda
; TITLE OF INVENTION: Novel Proteins and Nucleic Acids Encoding Same
; FILE REFERENCE: 21402-074 US
; CURRENT APPLICATION NUMBER: US/10/624,932
; CURRENT FILING DATE: 2003-07-21
; PRIOR APPLICATION NUMBER: 09/918,779
; PRIOR FILING DATE: 2001-07-03
; PRIOR APPLICATION NUMBER: 60/221,409
; PRIOR FILING DATE: 2000-07-28
; PRIOR APPLICATION NUMBER: 60/222,840
; PRIOR FILING DATE: 2000-08-04
; PRIOR APPLICATION NUMBER: 60/223,752
; PRIOR FILING DATE: 2000-08-08
; PRIOR APPLICATION NUMBER: 60/223,762
; PRIOR FILING DATE: 2000-08-08
; PRIOR APPLICATION NUMBER: 60/223,770
; PRIOR FILING DATE: 2000-08-08
; PRIOR APPLICATION NUMBER: 60/223,769
; PRIOR FILING DATE: 2000-08-08
; PRIOR APPLICATION NUMBER: 60/225,146
; PRIOR FILING DATE: 2000-08-14
```

```
; PRIOR APPLICATION NUMBER: 60/225,392
; PRIOR FILING DATE: 2000-08-15
; PRIOR APPLICATION NUMBER: 60/225,470
; PRIOR FILING DATE: 2000-08-15
; Remaining Prior Application data removed - See File Wrapper or PALM.
; NUMBER OF SEQ ID NOS: 61
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 2
; LENGTH: 898
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-624-932-2

Query Match 98.2%; Score 4698.5; DB 16; Length 898;
Best Local Similarity 98.7%; Pred. No. 0;
Matches 888; Conservative 2; Mismatches 7; Indels 3; Gaps 3;

QY 1 MAVRPGILWPAALLGIVLAAWLRGSGAQQSATVANPVPFANPDLLPHFLVEPEDIYIVKNKP 60
Db 1 MAVRPGILWPAALLGIVLAAWLRGSGAQQSATVANPVPFANPDLLPHFLVEPEDIYIVKNKP 60
QY 61 VLLVCKAVPATQIIPFKCNGEWVRQVDHVIERTSDGSSGEPTEMEVRINVSROQVEKVGLE 120
Db 61 VLLVCKAVPATQIIPFKCNGEWVRQVDHVIERTSDGSSGEPTEMEVRINVSROQVEKVGLE 120
QY 121 EYWCQCQVAMSSSCTTKSQKAYIRIARLKNFEOEPLAKEVSLBQGIPLCRPPEGIPPAE 180
Db 121 EYWCQCQVAMSSSCTTKSQKAYIRIARLKNFEOEPLAKEVSLBQGIPLCRPPEGIPPAE 180
QY 181 VEWLNRNEDLVDPDLDPNVYITREHSLVVRQARLADTANYTCVAKNIVARRRSASAAVIVY 240
Db 181 VEWLNRNEDLVDPDLDPNVYITREHSLVVRQARLADTANYTCVAKNIVARRRSASAAVIVY 240
QY 241 VNGGWSTWTEWSVCSASCGRWQKRSRSCNTPAPLNGGAFCEGQNVHDRTVSSLLVSDG 300
Db 241 VNGGWSTWTEWSVCSASCGRWQKRSRSCNTPAPLNGGAFCEGQNVHDRTVSSLLVSDG 300
QY 301 SWSPWKSWSACGLDCTHWRRECSDPAPRNGGECQGTDLDRNCTSDLCVHSASGPEV 360
Db 300 SWSPWKSWSACGLDCTHWRRECSDPAPRNGGECQGTDLDRNCTSDLCVHSASGPEV 359
QY 361 ALYVGLIAVAVCLVLLVLLVLYVYCRKKEGLSDVADSSILTSQFQPVSKPADNPHL 420
Db 360 ALYVGLIAVAVCLVLLVLLVLYVYCRKKEGLSDVADSSILTSQFQPVSKPADNPHL 419
QY 421 LTIQPDLS-TTTTYQGSCLCPQDGPSPKFQITNGHLLSLPLGGGRHTLHSSPTSEAEFF 479
Db 420 LTIQPDLS-TTTTYQGSCLCPQDGPSPKFQITNGHLLSLPLGGGRHTLHSSPTSEAEFF 479
QY 480 SRLSTQNYFRSLPRGTSNNMTYGTFFNLGGLMLPNTGISLILPPDAIPRGKIYIYLTLLH 539
Db 480 SRLSTQNYFRSLPRGTSNNMTYGTFFNLGGLMLPNTGISLILPPDAIPRGKIYIYLTLLH 539
QY 540 KPEDVRLPLAGCQTLSPVSCGPPGVLLTRPVILAMDHGCEPSDPSWSLRLKKQCEGS 599
Db 540 KPEDVRLPLAGCQTLSPVSCGPPGVLLTRPVILAMDHGCEPSDPSWSLRLKKQCEGS 599
QY 600 WEQDVLHLGEEAPSHLYYCOLEASACVFTFEQGRFALVGEALSVAAAKLKLFPAPVA 659
Db 600 WE-DVLHLGEEAPSHLYYCOLEASACVFTFEQGRFALVGEALSVAAAKLKLFPAPVA 658
QY 660 CTSLEYNIRVYCLDTHDALKEVVQLEKQGGOLIOBPRVLHFKDSVHNLRLSIHDPSS 719
Db 659 CTSLEYNIRVYCLDTHDALKEVVQLEKQGGOLIOBPRVLHFKDSVHNLRLSIHDPSS 718
QY 720 LKSKLLVSYQEIIPFFYHINWGTQRYLHCTFTLBRVSPSTSDACKLWVWQVEGDGQSF 779
Db 719 LKSKLLVSYQEIIPFFYHINWGTQRYLHCTFTLBRVSPSTSDACKLWVWQVEGDGQSF 778
QY 780 NFNITKDTFAELLALAESEAGVPALVGPSPAFKIPFLIRQKIISLDPCCRRGADWRTLAQ 839
Db 779 NFNITKDTFAELLALAESEAGVPALVGPSPAFKIPFLIRQKIISLDPCCRRGADWRTLAQ 838
```

QY 840 KLHLDLSLSPFASKPSPTAMILNLWEARHPNGNLSQLAAVAGLQPDAGLFTVSEAC 899
 DB 839 KLHLDLSLSPFASKPSPTAMILNLWEARHPNGNLSQLAAVAGLQPDAGLFTVSEAC 898

RESULT 4

US-09-933-261-5
 ; Sequence 5, Application US/09933261
 ; Publication No. US20030040046A1
 ; GENERAL INFORMATION:
 ; APPLICANT: Tessier-Lavigne, Marc
 ; Leonardo, E. David
 ; Hink, Lindsay
 ; Masu, Masayuki
 ; Kazuko, Keino-Masu
 ; TITLE OF INVENTION: Netrin Receptors
 ; NUMBER OF SEQUENCES: 8
 ; CORRESPONDENCE ADDRESSES:
 ; ADDRESSEE: SCIENCE & TECHNOLOGY LAW GROUP
 ; STREET: 268 BUSH STREET, SUITE 3200
 ; CITY: SAN FRANCISCO
 ; STATE: CALIFORNIA
 ; COUNTRY: USA
 ; ZIP: 94104
 ; COMPUTER READABLE FORM:
 ; MEDIUM TYPE: Floppy disk
 ; COMPUTER: IBM PC compatible
 ; OPERATING SYSTEM: PC-DOS/MS-DOS
 ; SOFTWARE: Patent In Release #1.0, Version #1.30
 ; CURRENT APPLICATION DATA:
 ; FILING DATE: 20-Aug-2001
 ; APPLICATION NUMBER: US/09/933,261
 ; CLASSIFICATION: <Unknown>
 ; PRIOR APPLICATION DATA:
 ; APPLICATION NUMBER: 08/808,982
 ; FILING DATE: <Unknown>
 ; ATTORNEY/AGENT INFORMATION:
 ; NAME: OSMAN, RICHARD A
 ; REGISTRATION NUMBER: 36,627
 ; REFERENCE/DOCKET NUMBER: UC96-217
 ; TELECOMMUNICATION INFORMATION:
 ; TELEPHONE: (415) 343-4341
 ; TELEFAX: (415) 343-4342
 ; INFORMATION FOR SEQ ID NO: 5:
 ; SEQUENCE CHARACTERISTICS:
 ; LENGTH: 898 amino acids
 ; TYPE: amino acid
 ; STRANDEDNESS: No. US20030040046A1 Relevant
 ; TOPOLOGY: No. US20030040046A1 Relevant
 ; MOLECULE TYPE: peptide
 ; SEQUENCE DESCRIPTION: SEQ ID NO: 5:

Query Match 95.0%; Score 4545.5; DB 10; Length 898;
 Best Local Similarity 94.7%; Pred. No. 0;
 Matches 852; Conservative 19; Mismatches 26; Indels 3; Gaps 3;

QY 1 MAVRPGIWPALLGIVLAALRGSGAQSATVANPVGANPDLLPHFLVEPEDVIVKNKP 60
 DB 1 MAVRPGIWPALLGIVLAALRGSGAQSATVANPVGANPDLLPHFLVEPEDVIVKNKP 60
 QY 61 VLLVCKAVPATQIFKCNKGEWVRQVDHVIERTSTGSGSEPTMEVIRINVSQQVEKVFGL 120
 DB 61 VLLVCKAVPATQIFKCNKGEWVRQVDHVIERTSTGSGSEPTMEVIRINVSQQVEKVFGL 120
 QY 121 EYWCQCVAWSSGTTKQKAYIRIARLKNFEPQPLAKEVSLQGVLPQCRPEGIPPAE 180
 DB 121 EYWCQCVAWSSGTTKQKAYIRIARLKNFEPQPLAKEVSLQGVLPQCRPEGIPPAE 180
 QY 181 VEWLRNEDLVDPSPDPNVYITREHSLVVRQARLADTANYTCVAKNIVARERSAAVIV 240
 DB 181 VEWLRNEDLVDPSPDPNVYITREHSLVVRQARLADTANYTCVAKNIVARERSAAVIV 240

QY 241 VNGWSTWTSVCSASCGRGWQKRSRSCCTNPAPINGGAFCEGQNVHRTVTSLLVSDG 300
 DB 241 VNGWSTWTSVCSASCGRGWQKRSRSCCTNPAPINGGAFCEGQNVHRTVTSLLVSDG 299
 QY 301 SWSWSKWSACGLDCTHWSRSCSDPAPRNGGEECGTDLTRNCTSDLCVHSASGPDV 360
 DB 300 SWSWSKWSACGLDCTHWSRSCSDPAPRNGGEECGTDLTRNCTSDLCVHSASGPDV 359
 QY 361 ALYVGLIAVAVCLVLLIIVLYCRKKEGLSDVADSSILTSFGPQPSVSKPSKADNPHL 420
 DB 360 ALYVGLIAVAVCLVLLIIVLYCRKKEGLSDVADSSILTSFGPQPSVSKPSKADNPHL 419
 QY 421 LTIQPDLS-TTTTQGSCLPRQDQSPKPFQLTNGHLLSPICGGRHHTLHSSPTSAEFPV 479
 DB 420 LTIQPDLS-TTTTQGSCLPRQDQSPKPFQLTNGHLLSPICGGRHHTLHSSPTSAEFPV 479
 QY 480 SRLSTQNYFRSLPRGTSMNTYGTNFGGLRLMIPNTGISLLIPDAIPRGKIYIYTLH 539
 DB 480 SRLSTQNYFRSLPRGTSMNTYGTNFGGLRLMIPNTGISLLIPDAIPRGKIYIYTLH 539
 QY 540 KPEDVRLPLAGCQTLSPVSCGPPGVLLTRPVILAMDHCGEPPSDMSLRLKKQSCBGS 599
 DB 540 KPEDVRLPLAGCQTLSPVSCGPPGVLLTRPVILAMDHCGEPPSDMSLRLKKQSCBGS 599
 QY 600 WEQDVTLHGEAPSHLYYCOLEASACVVFTEOLGRFALVGEALVAAKRLKLLFAPVA 659
 DB 600 WE-DVTLHGEAPSHLYYCOLEAGACVVFTEOLGRFALVGEALVAAKRLKLLFAPVA 658
 QY 660 CTSLEYNIRVYCLHDTHDAKEVVQLEKQGLIQRPVLHFKDSYHNLRSLSHDVPSS 719
 DB 659 CTSLEYNIRVYCLHDTHDAKEVVQLEKQGLIQRPVLHFKDSYHNLRSLSHDVPSS 718
 QY 720 LWSKLLVSYQEIPIFYHINWGTQRYLHCTFTLERVSPSTDLACKLWVQVGEQGSFSI 779
 DB 719 LWSKLLVSYQEIPIFYHINWGTQRYLHCTFTLERVSPSTDLACKLWVQVGEQGSFSI 778
 QY 780 NFNITKDTFAELLALAESEAGVPALVGFSAFKIPFLIROKIISSLDPPCRGADWRTIAQ 839
 DB 779 NFNITKDTFAELLALAESEAGVPALVGFSAFKIPFLIROKIISSLDPPCRGADWRTIAQ 838
 QY 840 KLHLDLSLSPFASKPSPTAMILNLWEARHPNGNLSQLAAVAGLQPDAGLFTVSEAC 899
 DB 839 KLHLDLSLSPFASKPSPTAMILNLWEARHPNGNLSQLAAVAGLQPDAGLFTVSEAC 898

RESULT 5

US-09-970-944-13
 ; Sequence 13, Application US/09970944
 ; Publication No. US20030204052A1
 ; GENERAL INFORMATION:
 ; APPLICANT: Heriman, John L
 ; APPLICANT: Rastelli, Luca
 ; APPLICANT: Shimkets, Richard A
 ; TITLE OF INVENTION: No. US20030204052A1
 ; TITLE OF INVENTION: Antibodies Directed Against these Proteins
 ; FILE REFERENCE: 21402-138
 ; CURRENT APPLICATION NUMBER: US/09/970,944
 ; CURRENT FILING DATE: 2002-05-02
 ; PRIOR APPLICATION NUMBER: 60/237,862
 ; PRIOR FILING DATE: 2000-10-04
 ; NUMBER OF SEQ ID NOS: 62
 ; SOFTWARE: Patent In Ver. 2.1
 ; SEQ ID NO 13
 ; LENGTH: 898
 ; TYPE: PRT
 ; ORGANISM: Rattus norvegicus
 ; US-09-970-944-13

Query Match 95.0%; Score 4545.5; DB 11; Length 898;
 Best Local Similarity 94.7%; Pred. No. 0;
 Matches 852; Conservative 19; Mismatches 26; Indels 3; Gaps 3;
 QY 1 MAVRPGIWPALLGIVLAALRGSGAQSATVANPVGANPDLLPHFLVEPEDVIVKNKP 60

```
Db 1 MAVRGLWFLVGLVLAAMLRGSGAQSSATVANPVGANPDLLPHLVEPEDEVYIVKNKP 60
QY 61 VLLVCKAVPATQIFPKCNGEWRQVDHVIERTSDGSGSEPTMEVRINVSRQVQKVFGL 120
Db 61 VLLVCKAVPATQIFPKCNGEWRQVDHVIERTSDGSGSEPTMEVRINVSRQVQKVFGL 120
QY 121 EYWCQCVAMSSGTTKSKAYIRIARLKNFQEPLEKEVSLEQGVLCRPEGIPPAE 180
Db 121 EYWCQCVAMSSGTTKSKAYIRIARLKNFQEPLEKEVSLEQGVLCRPEGIPPAE 180
QY 181 VEWLNRNEDLVDSLPNNVYITREHSLVVRQARLADTANYTCVAKNIVARRRSASAAVIV 240
Db 181 VEWLNRNEDLVDSLPNNVYITREHSLVVRQARLADTANYTCVAKNIVARRRSASAAVIV 240
QY 241 VNGGWSWTWESVCSASCGRWQKRSCTNPAPLNGGAFCEGQNV-QKTACATLCPVDG 300
Db 241 VNGGWSWTWESVCSASCGRWQKRSCTNPAPLNGGAFCEGQNV-QKTACATLCPVDG 300
QY 301 SWSWSKWSACGLDCTHWRSECSDPAPRNGECCGTDLDRNCTSDLCVHSASGPEV 360
Db 301 SWSWSKWSACGLDCTHWRSECSDPAPRNGECCGTDLDRNCTSDLCVHSASGPEV 360
QY 360 ALYVGLIAVAVCLFLLLLALGLIYCRKKEGLSDVDSSILTSFGFQPSIKPSKADNPHL 419
Db 360 ALYVGLIAVAVCLFLLLLALGLIYCRKKEGLSDVDSSILTSFGFQPSIKPSKADNPHL 419
QY 421 LTIQPDLS-TTITYQSGLCPRQDPSKPKFQLTNGHLLSPGGRHTLHSSPTSEAEFV 479
Db 421 LTIQPDLS-TTITYQSGLCPRQDPSKPKFQLTNGHLLSPGGRHTLHSSPTSEAEFV 479
QY 480 SRLSTQNYFRSLPRGTSNNYTFENFLGGRMLPNTGISLIPPDALPRKIYEIYTLH 539
Db 480 SRLSTQNYFRSLPRGTSNNYTFENFLGGRMLPNTGISLIPPDALPRKIYEIYTLH 539
QY 540 KPEDVRLPLAGCQTLSPVSVCGPPGVLLTRPVILAMDHCGEPSPDSWSLRKQKCEGS 599
Db 540 KPEDVRLPLAGCQTLSPVSVCGPPGVLLTRPVILAMDHCGEPSPDSWSLRKQKCEGS 599
QY 600 WQDVLHGEABRSHIYCOLEASACVTEQGLGRFALVGEALSVAAARLKLPPAPVA 659
Db 600 WQDVLHGEABRSHIYCOLEASACVTEQGLGRFALVGEALSVAAARLKLPPAPVA 659
QY 660 CTSLEYNIRVYCHLHDTHDALKEVQLEKQGLGQIPEPRVLFKDSYHNLRSIHDPVSS 719
Db 660 CTSLEYNIRVYCHLHDTHDALKEVQLEKQGLGQIPEPRVLFKDSYHNLRSIHDPVSS 719
QY 720 LWSKLLVSYQEIPIFYHNGTQRYLHCTFTLERSVSPSTDACKLWVWQVSGDGSFSI 779
Db 720 LWSKLLVSYQEIPIFYHNGTQRYLHCTFTLERSVSPSTDACKLWVWQVSGDGSFSI 779
QY 780 NFNIKTDTFAELLALLESAGVPALVGPAPKIPFLIROKIISLDPDPCRRGADWRTLAQ 839
Db 780 NFNIKTDTFAELLALLESAGVPALVGPAPKIPFLIROKIISLDPDPCRRGADWRTLAQ 839
QY 840 KLHDSHLSFFASKPSPPTAMINLWEARHPNGLNSQLAAVAGLQPDAGLFTVSEAC 899
Db 840 KLHDSHLSFFASKPSPPTAMINLWEARHPNGLNSQLAAVAGLQPDAGLFTVSEAC 899
```

RESULT 6

US-10-256-702-5

; Sequence 5, Application US/10256702

; Publication No. US20030059859A1

; GENERAL INFORMATION:

; APPLICANT: Tessier-Lavigne, Marc

; Leonardo, E. David

; Hink, Lindsay

; Masu, Masayuki

; Kazuko, Keino-Masu

; TITLE OF INVENTION: Netrin Receptors

; NUMBER OF SEQUENCES: 8

; CORRESPONDENCE ADDRESS:

```
; ADDRESS: SCIENCE & TECHNOLOGY LAW GROUP
; STREET: 268 BUSH STREET, SUITE 3200
; CITY: SAN FRANCISCO
; STATE: CALIFORNIA
; COUNTRY: USA
; ZIP: 94104
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/10/256,702
; FILING DATE: 27-Sep-2002
; CLASSIFICATION: <Unknown>
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US/09/933,261
; FILING DATE: 20-Aug-2001
; APPLICATION NUMBER: 08/808,982
; FILING DATE: <Unknown>
; ATTORNEY/AGENT INFORMATION:
; NAME: OSMAN, RICHARD A
; REGISTRATION NUMBER: 36,627
; REFERENCE/DOCKET NUMBER: UC96-217
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (415) 343-4341
; TELEFAX: (415) 343-4342
; INFORMATION FOR SEQ ID NO: 5:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 898 amino acids
; TYPE: amino acid
; STRANDEDNESS: No. US20030059859A1 Relevant
; TOPOLOGY: No. US20030059859A1 Relevant
; MOLECULE TYPE: peptide
; SEQUENCE DESCRIPTION: SEQ ID NO: 5:
; US-10-256-702-5
```

Query Match 95.0%; Score 4545.5; DB 14; Length 898;

Best Local Similarity 94.7%; Pred. No. 0;

Matches 852; Conservative 19; Mismatches 26; Indels 3; Gaps 3;

QY 1 MAVRGLWFLVGLVLAAMLRGSGAQSSATVANPVGANPDLLPHLVEPEDEVYIVKNKP 60

Db 1 MAVRGLWFLVGLVLAAMLRGSGAQSSATVANPVGANPDLLPHLVEPEDEVYIVKNKP 60

QY 61 VLLVCKAVPATQIFPKCNGEWRQVDHVIERTSDGSGSEPTMEVRINVSRQVQKVFGL 120

Db 61 VLLVCKAVPATQIFPKCNGEWRQVDHVIERTSDGSGSEPTMEVRINVSRQVQKVFGL 120

QY 121 EYWCQCVAMSSGTTKSKAYIRIARLKNFQEPLEKEVSLEQGVLCRPEGIPPAE 180

Db 121 EYWCQCVAMSSGTTKSKAYIRIARLKNFQEPLEKEVSLEQGVLCRPEGIPPAE 180

QY 181 VEWLNRNEDLVDSLPNNVYITREHSLVVRQARLADTANYTCVAKNIVARRRSASAAVIV 240

Db 181 VEWLNRNEDLVDSLPNNVYITREHSLVVRQARLADTANYTCVAKNIVARRRSASAAVIV 240

QY 241 VNGGWSWTWESVCSASCGRWQKRSCTNPAPLNGGAFCEGQNV-QKTACATLCPVDG 300

Db 241 VNGGWSWTWESVCSASCGRWQKRSCTNPAPLNGGAFCEGQNV-QKTACATLCPVDG 300

QY 301 SWSWSKWSACGLDCTHWRSECSDPAPRNGECCGTDLDRNCTSDLCVHSASGPEV 360

Db 301 SWSWSKWSACGLDCTHWRSECSDPAPRNGECCGTDLDRNCTSDLCVHSASGPEV 360

QY 361 ALYVGLIAVAVCLFLLLLALGLIYCRKKEGLSDVDSSILTSFGFQPSIKPSKADNPHL 420

Db 361 ALYVGLIAVAVCLFLLLLALGLIYCRKKEGLSDVDSSILTSFGFQPSIKPSKADNPHL 420

QY 421 LTIQPDLS-TTITYQSGLCPRQDPSKPKFQLTNGHLLSPGGRHTLHSSPTSEAEFV 479

Db 421 LTIQPDLS-TTITYQSGLCPRQDPSKPKFQLTNGHLLSPGGRHTLHSSPTSEAEFV 479


```
; SEQ ID NO 1
; LENGTH: 842
; TYPE: PRT
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: misc feature
; OTHER INFORMATION: Incyte ID No. US20040023244A1 6052371CD1
US-10-311-623-1

Query Match      91.5%; Score 4379.5; DB 16; Length 842;
Best Local Similarity 93.1%; Pred. No. 0;
Matches 838; Conservative 1; Mismatches 2; Indels 59; Gaps 3;

QY 1 MAVRPGLWPAIIIGVLAIAWLRGSGAQOQATVANVPVGPANPDLLPHFLVEPEDVYIVKNKP 60
DB 1 MAVRPGLWPAIIIGVLAIAWLRGSGAQOQATVANVPVGPANPDLLPHFLVEPEDVYIVKNKP 60
QY 61 VLLVCCKAVPATQIFPKCKGEWVRQVDHVIERTDSSGSEPTWEVRINVSROQVEKVFGL 120
DB 61 VLLVCCKAVPATQIFPKCKGEWVRQVDHVIERTDSSGSEPTWEVRINVSROQVEKVFGL 120
QY 121 EYWCQCVAMSSGGTTKQKAYIRIARLRKNFEQEPLAKEVSLQEGIVLPCRPEGIPPAE 180
DB 121 EYWCQCVAMSSGGTTKQKAYIRIARLRKNFEQEPLAKEVSLQEGIVLPCRPEGIPPAE 180
QY 181 VEWLNRNEDLVDPSPNVYITREHSLVVRQARLADTANYTCVAKNIVARRRSASAIVY 240
DB 181 VEWLNRNEDLVDPSPNVYITREHSLVVRQARLADTANYTCVAKNIVARRRSASAIVY 240
QY 241 VNGGWSWTWMSVCSACGRGWKRSRCTNPAPLNGGAFCEGQVHDRTVSSLLVSDG 300
DB 241 -----VDG 243

QY 301 SWSPKWSACGLDCTHWSRECDPAPRNGEECOGTDLDRNCTSDLCVHSASGPEDV 360
DB 244 SWSPKWSACGLDCTHWSRECDPAPRNGEECOGTDLDRNCTSDLCVHTASGPEDV 303
QY 361 ALYVGLIAVAVCLVLLVLLVLYCRKKEGLSDVADSSILTSQGPVSIKPSKADNPHL 420
DB 304 ALYVGLIAVAVCLVLLVLLVLYCRKKEGLSDVADSSILTSQGPVSIKPSKADNPHL 363
QY 421 LTIQPDLS-TTTTYQGSCLCPRODQSPKQFQNLNGHLLSPGGGRHTLHSSPTSABEFV 479
DB 364 LTIQPDLS-TTTTYQGSCLCPRODQSPKQFQNLNGHLLSPGGGRHTLHSSPTSABEFV 423
QY 480 SRLSTONYFRSLPRGTSNNMTYGTFFNLGGRMLPNTGISLLIPDPAIPRGKIYEIYLT 539
DB 424 SRLSTONYFRSLPRGTSNNMTYGTFFNLGGRMLPNTGISLLIPDPAIPRGKIYEIYLT 483
QY 540 KPEDVRLPLAGCQTLSPIVSCGPPGVLLTRPVILAMDHCGEPPSDSWSLRLKKQSCGS 599
DB 484 KPEDVRLPLAGCQTLSPIVSCGPPGVLLTRPVILAMDHCGEPPSDSWSLRLKKQSCGS 543
QY 600 WEQDVLHLGEEAPSHLYYQLEASACVYFTEQGRFALVGEALSVAALKLLFAPVA 659
DB 544 WE-DVLHLGEEAPSHLYYQLEASACVYFTEQGRFALVGEALSVAALKLLFAPVA 602
QY 660 CTSLEYNIRVYCLHDTHDALKEVVQLEKQGGQLIQEPRVLHFKDSYHNLRLSIHDVPS 719
DB 603 CTSLEYNIRVYCLHDTHDALKEVVQLEKQGGQLIQEPRVLHFKDSYHNLRLSIHDVPS 662
QY 720 LMKSKLLVSYOEIPFYHINWGTQRYLHCTFTLERSVSPSTDACKLWVWQVEGDGQSF 779
DB 663 LMKSKLLVSYOEIPFYHINWGTQRYLHCTFTLERSVSPSTDACKLWVWQVEGDGQSF 722
QY 780 NFNITKDTFAELALAESEAGVPALVGPFAFKIPFLIRQKIISLDPCCRRGADWRTLAQ 839
DB 723 NFNITKDTFAELALAESEAGVPALVGPFAFKIPFLIRQKIISLDPCCRRGADWRTLAQ 782
QY 840 KLHLDLSHLSFFASKPSPMTAMLNLEWHPFPNGNLSQAAAAVAGLQPDAGLFTVSEAC 899
DB 783 KLHLDLSHLSFFASKPSPMTAMLNLEWHPFPNGNLSQAAAAVAGLQPDAGLFTVSEAC 842
```

```
RESULT 9
US-09-970-944-14
; Sequence 14, Application US/09970944
; Publication No. US20030204052A1
; GENERAL INFORMATION:
; APPLICANT: Heriman, John L
; APPLICANT: Rastelli, Luca
; APPLICANT: Shmukets, Richard A
; TITLE OF INVENTION: No. US20030204052A1el Proteins and Nucleic Acids Encoding Same an
; FILE REFERENCE: 21402-138
; CURRENT APPLICATION NUMBER: US/09/970,944
; CURRENT FILING DATE: 2002-05-02
; PRIOR APPLICATION NUMBER: 60/237,862
; PRIOR FILING DATE: 2000-10-04
; NUMBER OF SEQ ID NOS: 62
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 14
; LENGTH: 544
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-970-944-14

Query Match      58.9%; Score 2819; DB 11; Length 544;
Best Local Similarity 99.6%; Pred. No. 4.2e-238;
Matches 540; Conservative 0; Mismatches 0; Indels 2; Gaps 2;

QY 359 DVALYVGLIAVAVCLVLLVLLVLYCRKKEGLSDVADSSILTSQGPVSIKPSKADNP 418
DB 4 DVALYVGLIAVAVCLVLLVLLVLYCRKKEGLSDVADSSILTSQGPVSIKPSKADNP 63
QY 419 HLLTITQPDLS-TTTTYQGSCLCPRODQSPKQFQNLNGHLLSPGGGRHTLHSSPTSABEF 477
DB 64 HLLTITQPDLS-TTTTYQGSCLCPRODQSPKQFQNLNGHLLSPGGGRHTLHSSPTSABEF 123
QY 478 FVSRSTQNYFRSLPRGTSNNMTYGTFFNLGGRMLPNTGISLLIPDPAIPRGKIYEIYLT 537
DB 124 FVSRSTQNYFRSLPRGTSNNMTYGTFFNLGGRMLPNTGISLLIPDPAIPRGKIYEIYLT 183
QY 538 LHKPEDVRLPLAGCQTLSPIVSCGPPGVLLTRPVILAMDHCGEPPSDSWSLRLKKQSC 597
DB 184 LHKPEDVRLPLAGCQTLSPIVSCGPPGVLLTRPVILAMDHCGEPPSDSWSLRLKKQSC 243
QY 598 GSWEQDVLHLGEEAPSHLYYQLEASACVYFTEQGRFALVGEALSVAALKLLFAP 657
DB 244 GSWE-DVLHLGEEAPSHLYYQLEASACVYFTEQGRFALVGEALSVAALKLLFAP 302
QY 658 VACTSLEYNIRVYCLHDTHDALKEVVQLEKQGGQLIQEPRVLHFKDSYHNLRLSIHDV 717
DB 303 VACTSLEYNIRVYCLHDTHDALKEVVQLEKQGGQLIQEPRVLHFKDSYHNLRLSIHDV 362
QY 718 SSLWSKLLVSYOEIPFYHINWGTQRYLHCTFTLERSVSPSTDACKLWVWQVEGDGQSF 777
DB 363 SSLWSKLLVSYOEIPFYHINWGTQRYLHCTFTLERSVSPSTDACKLWVWQVEGDGQSF 422
QY 778 SINFNITKDTFAELALAESEAGVPALVGPFAFKIPFLIRQKIISLDPCCRRGADWRT 837
DB 423 SINFNITKDTFAELALAESEAGVPALVGPFAFKIPFLIRQKIISLDPCCRRGADWRT 482
QY 838 AQKLDLSHLSFFASKPSPMTAMLNLEWHPFPNGNLSQAAAAVAGLQPDAGLFTVSEA 897
DB 483 AQKLDLSHLSFFASKPSPMTAMLNLEWHPFPNGNLSQAAAAVAGLQPDAGLFTVSEA 542
QY 898 EC 899
DB 543 EC 544

RESULT 10
US-09-933-261-6
; Sequence 6, Application US/09933261
; Publication No. US20030040046A1
```


QY 404 GFQVSIKPSKADNPHLLTIQPDLS-TTIVYQSLCPROQSPSPKFTNGHLLSPGLGG 462
Db 61 GFQVSIKPSKADNPHLLTIQPDLS-TTIVYQSLCPROQSPSPKFTNGHLLSPGLGG 120
QY 463 RHTLHSSPTSEAEFFVSRSTONYSFSLPRTGNTMYGTNEFLGGRLMIPNTGSLILIP 522
Db 121 RHTLHSSPTSEAEFFVSRSTONYSFSLPRTGNTMYGTNEFLGGRLMIPNTGSLILIP 180
QY 523 PDAIPRGKIYIYITLHKPEVDRLPLAGCQTLSPIVSCGPPGVLLTRPVILAMDHCGEP 582
Db 181 PDAIPRGKIYIYITLHKPEVDRLPLAGCQTLSPIVSCGPPGVLLTRPVILAMDHCGEP 240
QY 583 SPDWSLRLKQSCGSEWQDVLHGEAPSHLYYQLEASACVYFTQOLGRFALVGEAL 642
Db 241 SPDWSLRLKQSCGSEWQDVLHGEAPSHLYYQLEASACVYFTQOLGRFALVGEAL 299
QY 643 SVAARLKLILFAPVACTSLEYNIRVYCLDTHDALKVVOLEKQGLGQIQLIEPRVLHF 702
Db 300 SVAARLKLILFAPVACTSLEYNIRVYCLDTHDALKVVOLEKQGLGQIQLIEPRVLHF 359
QY 703 XDSYHNLRLSHDVPSLWKSLLVSYQEIPIFYHIWNGTQRYLHCTFTLERSVSPSTDLA 762
Db 360 XDSYHNLRLSHDVPSLWKSLLVSYQEIPIFYHIWNGTQRYLHCTFTLERSVSPSTDLA 419
QY 763 CKLWVQVEGQSQSFNFINITKTRFAELLALAESEAGVPALVGPSAFKIPFLIRQKIIS 822
Db 420 CKLWVQVEGQSQSFNFINITKTRFAELLALAESEAGVPALVGPSAFKIPFLIRQKIIS 479
QY 823 SLDPCCRGADWRTLAQKLHLDLSHLSFASKPSPTAMILNWEARHFNGLSQAALAAVA 882
Db 480 SLDPCCRGADWRTLAQKLHLDLSHLSFASKPSPTAMILNWEARHFNGLSQAALAAVA 539
QY 883 GLGQPDAGLFT-VSEAE 899
Db 540 GTXPAGRWLLSQCSAEAC 557
RESULT 12
US-09-970-944-15
; Sequence 15, Application US/09970944
; Publication No. US20030204052A1
; GENERAL INFORMATION:
; APPLICANT: Herrman, John L
; APPLICANT: Rastelli, Luca
; APPLICANT: Shinkets, Richard A
; TITLE OF INVENTION: No. US20030204052A1el Proteins and Nucleic Acids Encoding Same an
; TITLE OF INVENTION: Antibodies Directed Against these Proteins
; FILE REFERENCE: 21402-138
; CURRENT APPLICATION NUMBER: US/09/970,944
; CURRENT FILING DATE: 2002-05-02
; PRIOR APPLICATION NUMBER: 60/237,862
; PRIOR FILING DATE: 2000-10-04
; NUMBER OF SEQ ID NOS: 62
; SOFTWARE: Patentin Ver. 2.1
; SEQ ID NO 15
; LENGTH: 931
; TYPE: PRT
; ORGANISM: Caenorhabditis elegans
US-09-970-944-15
Query Match 56.5%; Score 2705.5; DB 11; Length 931;
Best Local Similarity 56.4%; Pred. No. 9e-228;
Matches 515; Conservative 153; Mismatches 214; Indels 31; Gaps 12;
QY 9 PALLGIVLAWLGRSGNAQQA---TVANPYGANPDLLPHFLVEPEDEVYIVKKNPVLLVC 65
Db 26 PAL--ALLSASGTGSAQADDEFFHELPETFPSPPEPLPHLEPEEAYIVKKNPNLYC 83
QY 66 KAVPATQIFKCNKGEWVRQVDHVIERTSDGSGEPTMEVNRVSRQVKEVFGLEEYWCQ 125
Db 84 KASPATQIFKCNSEWVHQKHVVDERVDETSGLIVREVSIRQVVEELFGEDYWCQ 143

QY 126 CVAWSSGTTKSKAYIRIARLNKFNFOEPLAKEVSLEQGIIVLPCRPPEGIPPAEVEWL 185
Db 144 CVAWSSGTTKSKAYIRIAYLRKTFEOPLEKVSLEQEVLLQCRPEGIPPAEVEWLK 203
QY 186 NEDLVDPSPDPNVYITREHSLVVRQARLADTANTYCVAKNIVARRRSASAAVIVYVNGW 245
Db 204 NEDIIDPAEDRNFYITIDHNLIIKQARLSDTANTYCVAKNIVAKRKSTTATVIVYVNGW 263
QY 246 STWTEWSVCSACGRGQKSRSCNTNAPLNGGAFCEGQNVHDRTVSSLLVSDGSPW 305
Db 264 STWTEWSVCSACGRGQKSRSCNTNAPLNGGAFCEGQSV-QKIACITLCPVDGRWTSW 322
QY 306 SKWACGLDCHWMSRECSDPAPRNGEECGTDLDTNCTSDLCVHSASGPEVDVALYVG 365
Db 323 SKWTCGTCTHWRRECTAPAKNGGKCDGLVQSKNCTDGLCMAQAPSDDDVALYVG 382
QY 366 L-IAVAVCLVLLVLLVYCRKKEGLSDVADSSILTSFGQPVSIKPSKADNPHLLTIQ 424
Db 383 IVIATVCLAITVVVALFVYKRNHRDPESDIIDSSALNGGQFPVNIKAARQD---LLAVP 439
QY 425 PDL-STITTYQSGSLCPROQSPKFTNGHLLSPGLGGRLTHHSS-----PTSEAEFFV 479
Db 440 PDLTSAAMYRGVPVALHD-VSDKIPMTNPSILDPNLKIKVYNSSGAVTPQDDLAEP 498
QY 480 SRLS---TONYF-----RSLPRGT--SNMTYGTENFLGGRLMIPNTGSLILIP 525
Db 499 SKLSPQMTQSLLENALNKNQSLARQTDPSCTAGTFENSLGGHLIIPNSGVSLIIPAGA 558
QY 526 IPRGKIYIYITLHKPEVDRLPLAGCQTLSPIVSCGPPGVLLTRPVILAMDHCGEPSPD 585
Db 559 IPOGRVYEMVTVHRKKNRPPMEDSQTLTPVWSCGPPGALLTRPVILTLLHHCADPSTE 618
QY 586 SWSLRLKQSCGSEWQDVLHGEAPSHLYYQLEASACVYFTQOLGRFALVGEALSAVA 645
Db 619 DMKIQLKNQAVQGWQW-DVVVVVEENFTTPCYIQLDABACHILTENLSTYALVGQSTYKA 677
QY 646 AAKRLKLLFAPVACTSLEYNIRVYCLDTHDALKVVOLEKQGLGQIQLIEPRVLHF 705
Db 678 AAKRLKLAIFGPLCCSSLEYSIRVYCLDDTDALKEVLQERQMGQQLLEEPKALHFKGS 737
QY 706 YHNLRLSHDVPSLWKSLLVSYQEIPIFYHIWNGTQRYLHCTFTLERSVSPSTDIA 765
Db 738 IHNRLSLHDIHSLWKSLLVSYQEIPIFYHIWNGTQRYLHCTFTLERSVSPSTDIA 797
QY 766 WTVQVEGQSQSFNFINITKTRFAELLALAESEAGVPALVGPSAPKIPFLIRQKIIS 825
Db 798 CYRQVEGEQIFQLNCTVSEETPGIDPLDLPASTITTVTGPSAFSIPILPIKQLCSSLD 857
QY 826 PPCRGADWRTLAQKLHLDLSHLSFASKPSPTAMILNWEARHFNGLSQAALAAVA 885
Db 858 APQTRGHDRWRLAHKLNLDRLYNYFATKSSPTGVILDLWEAQNFDPGNLSMLAAVLEEW 917
QY 886 QPDAGLFTVSEAB 898
Db 918 RHETVTVSLAAEQ 930
RESULT 13
US-10-087-684-35
; Sequence 35, Application US/10087684
; Publication No. US20040029116A1
; GENERAL INFORMATION:
; APPLICANT: Edinger, Shlomit R.
; APPLICANT: MacDougall, John R.
; APPLICANT: Millet, Isabelle
; APPLICANT: Miller, Karen
; APPLICANT: Stone, David J.
; APPLICANT: Grosse, William M.
; APPLICANT: Lepley, Denise M.
; APPLICANT: Rieger, Daniel K.
; APPLICANT: Burgess, Catherine E.
; APPLICANT: Casman, Stacie, J.
; APPLICANT: Sptek, Kimberly A.

;	Prior Filing Date:	2000-10-06	
;	Prior Application Number:	60/238,379	
;	Prior Filing Date:	2000-10-06	
;	Prior Application Number:	60/238,402	
;	Prior Filing Date:	2000-10-06	
;	Prior Application Number:	60/238,384	
;	Prior Filing Date:	2000-10-06	
;	Prior Application Number:	60/238,373	
;	Prior Filing Date:	2000-10-06	
;	Prior Application Number:	60/238,372	
;	Prior Filing Date:	2000-10-06	
;	Prior Application Number:	60/238,383	
;	Prior Filing Date:	2000-10-06	
;	Prior Application Number:	60/275,892	
;	Prior Filing Date:	2001-03-14	
;	Prior Application Number:	60/296,860	
;	Prior Filing Date:	2001-06-08	
;	Number of Seq ID NOS:	198	
;	Software:	Patentin Ver. 2.1	
;	Seq ID NO	121	
;	Length:	931	
;	Type:	PRT	
;	Organism:	Mus musculus	
;	US-09-972-211-121		
;	Query Match	56.5%; Score 2705.5; DB 12; Length 931;	
;	Best Local Similarity	56.4%; Pred. No. 9e-228;	
;	Matches	535; Conservative 153; Mismatches 214; Indels 31; Gaps 12;	
QY	9	PALLGVLAAMLRGSGAQAQSA---TVANPVPGANPDLLPHFLVPEPDIYVKNKPVLLVC	65
DB	26	PAL--ALLASGCTGAAQDEFFHELPEPDPPEPLPHELPEEYAVIYVKNKPVNYC	83
QY	66	KAVPATQIEFKNGEWRQVDHVIERSDGSGETMVRINVSROQVEKVFGLLEEWYCQ	125
DB	84	KASPATQIYFKCNSEWVQKHVDVDSGLTVREVSIEISROQVEELFGPDEYWCQ	143
QY	126	CVAMSSGGTYSKAYIRIARLKNFEOEPLAKEVSLQGIKVLPCRPPEGIPPAEVEWLR	185
DB	144	CVAMSSAGTTKSKAYVRYAYLRKTFEQEPLGKEVSLQEVLLQCRPEGPVPAEVEWLK	203
QY	186	NEDVDPSLDPNVYITREHSLVVRQARLADTANYTCVAKNIVARRRSASAAVIVYVNGW	245
DB	204	NEDIIDPAEDRNFYITIDHNLIIKQARLSDTANYTCVAKNIVAKRKSTTATVIVYVNGW	263
QY	246	STWTEWSVCSACGRGWKRSCTNPAPLNGACFCEQNVHDTVSVSLLSVVDGWSWP	305
DB	264	STWTEWSVCSNCRGQYQKRTCTNPAPLNGACFCEQSV-QKIACITLCPVDGRWTSW	322
QY	306	SKWSACGLDCTHWRSGESDPAIPRGGBECQGTDLDTNRCTSDLCVHSASGPEDVALYVG	365
DB	323	SKWSTGCTEHTHRRRECTAPAPKNGKDCGLVLQSNCTDGLCMQAPDSDVALYVG	382
QY	366	L-IAVAVCVLHLLLVILVYCKEGLSDVDASSILTSFGPVSIRKPSKADNPHLLTIQ	424
DB	383	IVIAVTVCLATVWVALFVYRKNHRDFESDIIDSSALNGGFPQVNIKAARD---LLAVP	439
QY	425	PDL-STITTYQCSLCPRGQSPKQLNGLHLLSPLGGRHTLHHS-----PTSEAEFV	479
DB	440	PDLTSAAMYRGFPVVALHD-VSDKIPMTNSPLDPLNKLKIVYNSSGAVTPQDDLAFFS	498
QY	480	SRLS-----TQNYF-----RSLPRGT--SNMTYGTENFLGRLMIPINTGISLILIPDA	525
DB	499	SKLSPQMTOSLLENEALNKNQSLARQDTPCTAGTFNSLGGHLLTIPNSGVSLIIPAGA	559
QY	526	IPRGIYEYTLTHKPEDVRLPLAGQTLSPISVSGPPGVLLTRPVLAMDMHCHGEPSPD	585
DB	559	IPQGRVYEMVYVHRKENRPMWDSQTLTPVVSFGPGALLTFPVLTLHHCADPSTE	618
QY	586	WSLRLKQSCGSGEWDVLLHGEBAPSHLYYCOLEASACVYFTQLGRFVALGEALSV	645

GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - protein search, using sw model

Run on: October 4, 2004, 18:23:31 ; Search time 50 seconds
(without alignments)
1729.524 Million cell updates/sec

Title: US-09-970-944-2
Perfect score: 4787
Sequence: 1 MAVRGLMPALLGIVLAWL.....AVAGLQPDAGLFTVSEAC 899
Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 283366 seqs, 96191526 residues
Total number of hits satisfying chosen parameters: 283366

Minimum DB seq length: 0
Maximum DB seq length: 2000000000
Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : PIR 78.*
1: pir1.*
2: pir2.*
3: pir3.*
4: pir4.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	950.5	19.9	919	2 T32541	unc-5 protein - Ca
2	950.5	19.9	947	1 B44294	unc-5 protein, lon
3	280	5.8	1584	2 T00026	brain-specific ang
4	273.5	5.7	1172	1 TSHU2	thrombospondin 2 p
5	273.5	5.7	1172	2 A42587	thrombospondin 2 p
6	270.5	5.7	1074	2 JC5928	semaphorin F precu
7	267.5	5.6	984	2 T00326	hypothetical prote
8	267.5	5.6	1522	2 T00028	brain-specific ang
9	263.5	5.5	1444	2 T18856	angiogenesis inhib
10	258	5.4	1572	2 T00027	brain-specific ang
11	248	5.2	1170	2 A40538	thrombospondin 1 p
12	246	5.1	1170	1 TSHU1	thrombospondin 1 p
13	243.5	5.1	1178	1 A39804	thrombospondin 1 p
14	220.5	4.6	1651	2 T14160	thrombospondin pre
15	217.5	4.5	1612	2 T30805	transmembrane rece
16	211.5	4.4	1265	1 A37967	cutti protein - mo
17	208.5	4.4	469	1 S29126	neural cell adhesi
18	206.5	4.3	254	2 T15952	properdin precursu
19	202.5	4.2	437	2 S05478	properdin - mouse
20	198.5	4.1	788	2 T25061	hypothetical prote
21	188	3.9	423	2 T29549	hypothetical prote
22	185	3.9	1344	2 T14316	hypothetical prote
23	184	3.8	957	2 T15976	rig-1 protein - mo
24	183.5	3.8	1273	2 T42405	hypothetical prote
25	183.5	3.8	1863	2 S46217	sax-3 protein - Ca
26	176.5	3.7	860	2 T16892	protein-tyrosine-p
27	175.5	3.7	1736	2 A4747	hypothetical prote
28	170.5	3.6	1745	2 A46431	tight junction pro
29	169	3.5	934	1 A34372	complement C6 prec

30	167.5	3.5	1907	2 S50893	protein-tyrosine-p
31	166	3.5	837	2 T00355	hypothetical prote
32	160.5	3.4	550	2 T47158	hypothetical prote
33	152.5	3.2	1499	2 I50212	protein-tyrosine-p
34	152	3.2	951	2 T00017	gene ADAMTS-1 prot
35	152	3.2	1501	2 I58148	protein-tyrosine-p
36	151	3.2	2165	2 T21371	hypothetical prote
37	150	3.1	1277	2 T30532	neural cell adhesi
38	149	3.1	654	2 T23247	hypothetical prote
39	148.5	3.1	946	1 A47299	ror-related recept
40	148	3.1	1091	2 A58532	glial cell membran
41	147	3.1	843	1 A27340	complement C7 prec
42	147	3.1	888	2 S23065	ufo protein - mous
43	146.5	3.1	584	1 C8HUA	kinase-like protei
44	146	3.0	1051	2 A39712	protein-tyrosine-p
45	145.5	3.0	2029	1 TDFFLK	

ALIGNMENTS

RESULT 1

T32541
unc-5 protein - Caenorhabditis elegans
C:Species: Caenorhabditis elegans
C:Date: 29-Oct-1999 #sequence_revision 29-Oct-1999 #text_change 31-Jan-2000
R:Accession: T32541
R:Ratrefille, P.
submitted to the EMBL Data Library, December 1997
A:Description: The sequence of C. elegans cosmid B0273.
A:Reference number: Z21187
A:Accession: T32541
A>Status: preliminary; translated from GB/EMBL/DBJ
A:Molecule type: DNA
A:Residues: 1-919 <LAT>
A:Cross-references: EMBL:AF036698; PIDN:AAB86355.1; GSPDB:GN00022; CESP:B0273.4a
A:Experimental source: strain Bristol N2; clone B0273
C:Genetics:
A:Gene: unc-5; CESP:B0273.4a
A:Map position: 4
A:Introns: 41/3; 108/1; 142/3; 201/1; 323/2; 553/1; 859/3
C:Superfamily: unc-5 protein; immunoglobulin homology; SH3 homology; thrombospondin type

Query Match	19.9%	Score	950.5	DB 2	Length	919;
Best Local Similarity	28.5%	Pred. No.	1.6e-61;			
Matches	265;	Conservative	165;	Mismatches	375;	Indels 125; Gaps 31;
QY	49	EPEDVYTKNKPVLVCKAVPATQIFKCNCEWVRQVDHVIET--STDGSSGEPTMEVRI	106			
Db	9	QPKSGYVIRNKPRLQCRANHATKIRYKCSKWID--DSRIEKLIGTDSTSGVGVIDASV	66			
QY	107	NVSRQQVEKVFGLBEYWCQCQVAVSSSG-----TTKSQKAVIRIARLNKPFCEPLAKEVS	161			
Db	67	DISRIDVDTSQVDAFQCQYA---SGDDQDVVASDVATVHLAYMKHFLKSPVQRVQ	123			
QY	162	LEQGIPLPCPPGIPPAEWEIWNEDLDVPSLDPNVYITREHSLVVVRQARLADTANYTC	221			
Db	124	EGTTLQPCAPESDPKAEITWKDGVVQVP--DANVIRASDGLSIWSAARLSDSGNYTC	181			
QY	222	VAKNIVARRRSASAIVVYVGGWSTWETW-SVCSASCG-----RGWQKR	265			
Db	182	EATNVANSRTDPVEVOIYVDGGSWSPWIGTCHVDCPLLRQHAHRIRDPHDPHQRR	241			
QY	266	SRSCITNPALINGAFCEGQNVHRTVSSLLVSDGWSMWSKWSACGLDCTHWSRECS	325			
Db	242	TRICNNPAPLNDGEYCKGEBEMTRS-CKVCKLDGGSSWSWSACSSCHRRVTRACTV	300			
QY	326	PAPNGECCQGTDLTRNCTSDLCVHSAAG--PEDVALYVGLIAVAVCLVLLVLLV	383			
Db	301	PPPMNGGQPCFGDMLTQECPAQLCTADSSRIVISDTAVYGSVASIFIVASFILAILAMF	360			
QY	384	YCR-----KKEGLSDVADS-----SILTSGFQF	407			

Db 361 CCKGNSKKSLPKLPQKNSEKAGIYSSPPGVRLLLEHQHQTLLGEKISSCSQYFEP 420
 Qy 408 VSIKPSKADNPILLTIQPDLSSTTTTYQGLSCPRQDGPSPKPOLNWHLLSPGLGGRTTLH 467
 Db 421 PPL-----PHSTTLR---SGKSAFSGYSTRNAGSRAALIQECSSSSSGGKRTMLR 470
 Qy 468 HSPST-SEAEFVSRSLSTONYFRSLPRGTS-NMTYGTFFNLGRLMIPNTGISLILPDA 525
 Db 471 TSSNSCDDDDNYATLYDYMEDKSVLGLDTSQNIIVAAQIDNSGARLSKSGARLIVPELA 530
 Qy 526 IPRGKIYEIYTLHKPEDVRLPLAGCOTLLSPIVSCGPPGV-----LLTRPVILAMDH 578
 Db 531 VEKEM--LYLAVSDTLTDQPHLKPIESALSPVIVIGQCDVMSAHNDILRRPVVVSFRH 588
 Qy 579 CGEFP-DSWSLRLLKQSCGSEWQDVLHLEAPSHLYYCOLEASA-----CYVFT 629
 Db 589 CASTFPRDNWQFTL--YADEGSGWQKAVTIGENLNTNMVQFQPGKKNDFGFWCHVMT 646
 Qy 630 EQLGFPALVGEAL--SVAAAKRLKLLLPAPVACTSLE--YNIRVYCLHDTHDALKEVVQL 685
 Db 647 YSLARLMLAGHPRRNSLSAAKRVHLAVFGPTMSAYRRPFELRVYCVPTGAAMESVWKQ 706
 Qy 686 EQLGGQLIQEPR--VLHFKDSYHNRLSIHDV-PSSLWKSLLVSYQIPIFYHIWNGTQ 742
 Db 707 ED--GSRLLCESNDFILNEKG---NLICIEDVIFGFCGPEVVEISETQHRFV--AQ 758
 Qy 743 RYLHCTFTLERSVPSSTDLACKLWVQVEGQGSFINSFNITKOTRFAELIALESEAGVP 802
 Db 759 NGLHCSLKFPRKEINGSQFSTRVIVYQKASSTPEVMV--EVSNEPELYDATSEEREKGSV 816
 Qy 803 ALVGSAFKIPLFLPKIOLSISSLDPPRCGADWRTIAQLHLDLSHLSFPASKP--SPTAMI 860
 Db 817 CV----EFLRPGVMDLRLDMPNESHSDWRGLAKLHYDRILQFASFPDCSPTSLL 872
 Qy 861 LNLWEARHPNGN-LSQAAAAVAGLQCPDA 889
 Db 873 LDLWEASSSGSARAVPDLLOTLRVWGRPD 902

RESULT 2
 B44294
 unc-5 protein, long form - Caenorhabditis elegans
 N:Contains: unc-5 protein, short form
 C:Species: Caenorhabditis elegans
 C>Date: 30-Apr-1993 #sequence revision 28-Jul-1995 #text_change 05-Nov-1999
 C:Accession: B44294; T32540; A44294
 R:Leung-Hageteijn, C.; Spence, A.M.; Stern, B.D.; Zhou, Y.; Su, M.W.; Hedgecock, E.M.; Cell 71, 289-299, 1992
 A:Title: UNC-5, a transmembrane protein with immunoglobulin and thrombospondin type 1 do
 A:Reference number: A44294; PMID:93046629; PMID:1384987
 A:Contents: variety Bergerac
 A:Accession: B44294
 A:Molecule type: DNA
 A:Residues: 1-947 <LEU>
 A:Cross-references: GB:S47168; NID:G258527; PIDN:AAB21867.1; PID:G258529
 A>Note: sequence extracted from NCBI backbone (NCBIN:116668, NCBIN:116670, NCBIN:116672,
 A>Note: authors translated the codon CTA for residue 642 as val; sequence shown follows
 A>Note: mRNA lacking the first exon is equally prevalent
 R:Latreille, P.
 Submitted to the EMBL Data Library, December 1997
 A:Description: The sequence of C. elegans cosmid B0273.
 A:Reference number: Z21187
 A:Accession: T32540
 A>Status: preliminary; translated from GB/EMBL/DDBJ
 A:Molecule type: DNA
 A:Residues: 1-947 <LAT>
 A:Cross-references: EMBL:AF036698; PIDN:AAB88356.1; GSPDB:GN00022; CESP:B0273.4b
 A:Experimental source: strain Bristol N2; clone B0273
 C:Genetics:
 A:Gene: unc-5
 A:Map position: 4
 A:Introns: 28/1; 69/3; 136/1; 170/3; 229/1; 351/2; 581/1; 886/3
 C:Function:

A:Description: required for guidance of pioneering axons and cells migrating dorsally along
 C:Superfamily: unc-5 protein; immunoglobulin homology; SH3 homology; thrombospondin type
 C:Keywords: alternative splicing; duplication; glycoprotein; receptor; transmembrane protein
 F:30-947/Product: unc-5 protein, short form #status predicted <ALT>
 F:46-116/Domain: immunoglobulin homology <IM2>
 F:153-211/Domain: immunoglobulin homology <IM2>
 F:229-300/Domain: thrombospondin type 1 repeat homology #status atypical <THR1>
 F:301-354/Domain: thrombospondin type 1 repeat homology <THR2>
 F:365-390/Domain: transmembrane #status predicted <TM>
 F:512-559/Domain: SH3 homology <SH3>
 F:53-114,65-112,160-209/Dissulfide bonds: #status predicted
 F:206/Binding site: carbohydrate (Asn) (covalent) #status predicted

Query Match 19.9%; Score 950.5; DB 1; Length 947;

Best Local Similarity 28.5%; Pred. No. 1.7e-61;

Matches 265; Conservative 165; Mismatches 375; Indels 125; Gaps 31;

Qy 49 EPEDVYVKNKPVLLVCKAVPATQIFKCNCEWVRQVDHVIET--STGSGSEPTMEVRI 106

Db 37 QPKSGYVIRNKLRLQCRANHATKIRYKCSKWID--DSRIEKLIGTSTGSGVYDASV 94

Qy 107 NVSRQOQEVKGLBEYWCOCVAMSSG-----TTKQKAYIRIARLRKNFQEPLEKVEVS 161

Db 95 DISRIDVDTSGHVDVAFQOCYA--SGDDDDQDVASDVATVHLAYMRKHLKSPVQRVQ 151

Qy 162 LEQGIPLPCRPEGIPPAEVEWLRNEDLVDPSPDENVYITREHSLVVRQARLADTANYTC 221

Db 152 EGTTLQLPCQAPSDPKAELTWYKGVVQV--DANVRASDGLSMAARLSDSGNYTC 209

Qy 222 VAKNTVARRRSASAAVIVYNGSGWTWEM-SVCASCG-----RGWQKR 265

Db 210 EATNVANGKTKDPVEVQIYVDGWSWSPWIGTCHVDCLLRQHAHRIRDPHDLVPHQR 269

Qy 266 SRSCNTPAPLNGARCEGQNVHDTVSSILVSDGWSWSPKAGLDCTHWRGECSD 325

Db 270 TETCNAPLNDGEYCKGEEEMTR8-CKVPCKLGGWSWSWDSACSSCHRYRFACTV 328

Qy 326 PAPRNGEGECQGTDLTDRNCTSDLCVHSGASG--PEDVALYVGLIAVAVCLVLLLVLLIV 383

Db 329 PPPMNGGQCFGDDLTQCEPAQLCTADSSRIVISDTAVYGSVASIFIVASFIATLAMP 388

Qy 384 YCR-----KXEGLDSDVADS-----SILTSGFQ 407

Db 389 CCKRGNKSKSLPKLPQKNSEKAGIYSSPPGVRLLLEHQHQTLLGEKISSCSQYFEP 448

Qy 408 VSIKPSKADNPILLTIQPDLSSTTTTYQGLSCPRQDGPSPKPOLNWHLLSPGLGGRTTLH 467

Db 449 PPL-----PHSTTLR---SGKSAFSGYSTRNAGSRAALIQECSSSSSGGKRTMLR 498

Qy 468 HSPST-SEAEFVSRSLSTONYFRSLPRGTS-NMTYGTFFNLGRLMIPNTGISLILPDA 525

Db 499 TSSNSCDDDDNYATLYDYMEDKSVLGLDTSQNIIVAAQIDNSGARLSKSGARLIVPELA 558

Qy 526 IPRGKIYEIYTLHKPEDVRLPLAGCOTLLSPIVSCGPPGV-----LLTRPVILAMDH 578

Db 559 VEKEM--LYLAVSDTLTDQPHLKPIESALSPVIVIGQCDVMSAHNDILRRPVVVSFRH 616

Qy 579 CGEFP-DSWSLRLLKQSCGSEWQDVLHLEAPSHLYYCOLEASA-----CYVFT 629

Db 617 CASTFPRDNWQFTL--YADEGSGWQKAVTIGENLNTNMVQFQPGKKNDFGFWCHVMT 674

Qy 630 EQLGFPALVGEAL--SVAAAKRLKLLLPAPVACTSLE--YNIRVYCLHDTHDALKEVVQL 685

Db 675 YSLARLMLAGHPRRNSLSAAKRVHLAVFGPTMSAYRRPFELRVYCVPTGAAMESVWKQ 734

Qy 686 EQLGGQLIQEPR--VLHFKDSYHNRLSIHDV-PSSLWKSLLVSYQIPIFYHIWNGTQ 742

Db 735 ED--GSRLLCESNDFILNEKG---NLICIEDVIFGFCGPEVVEISETQHRFV--AQ 786

Qy 743 RYLHCTFTLERSVPSSTDLACKLWVQVEGQGSFINSFNITKOTRFAELIALESEAGVP 802

Db 787 NGLHCSLKFPRKEINGSQFSTRVIVYQKASSTPEVMV--EVSNEPELYDATSEEREKGSV 844

Qy	803	ALVGP	SAFKI	PELIR	QKLI	SISSLD	PPCR	CGAD	NR	TT	IAQK	LHLS	HSHL	FF	SKP	---SPTAM	860
Db	845	CV----	EPRL	PGV	KDEL	ARLL	MDN	ESH	S	---	---	---	---	---	---	---	900
Qy	861	LNL	WEAR	HF	PNGN	---LS	QAAA	AVAG	LQ	PD	A	889					
Db	901	LDL	WEAS	SS	SGS	AR	AV	PD	LL	QTL	LR	VN	GR	PD	A	930	

RESULT 3

T00026
brain-specific angiogenesis inhibitor 1 - human
N;Alternate names: BAIL protein
C;Species: Homo sapiens (man)
C;Date: 22-Jan-1999 #sequence_revision 22-Jan-1999 #text_change 12-Feb-1999
C;Accession: T00026
R;Nishimori, H.; Shiratsuchi, T.; Urano, T.; Kimura, Y.; Kiyono, K.; Tatsumi, K.; Yoshida
submitted to the EMBL Data Library, June 1997
A;Reference number: Z14064
A;Accession: T00026
A;Status: translated from GB/EMBL/DDBJ
A;Molecule type: mRNA
A;Residues: 1-1584 <NIS>
A;Cross-references: EMBL:AB005297; NID:d1175078; PID:d1024528
A;Experimental source: brain

```

C;Genetics:
A;Gene: GDB:BA11
A;Cross-references: GDB:9838088; OMIM:602682
A;Map position: 8q24-8q24
C;Superfamily: thrombospondin type 1 repeat homology
F;408-462/Domain: thrombospondin type 1 repeat homology <THR3>

Query Match          5.8%; Score 280; DB 2; Length 1584;
Best Local Similarity 31.8%; Pred.No. 6.5e-12;
Matches 74; Conservative 38; Mismatches 93; Indels 28; Gaps 10;

QY      124  CQCVAMSSSGTTKSQAKYIARIARKNFEOEPLAKEVLSLEGIVLPCRPPGIGIPAEVW 183
Db      309  CNREACGPAGRTSRSSOSLSTDARR---REELGDEL---QQGFPPA-PQTGDPAAE-EW 360
QY      184  LRNEDLVDPSPDNVYITREHSLVVRQARLADTANYTCVAKNIIVARRSSAAAVIVVNG 243
Db      361  --SPWVSVCSTCGEGWQTR-----TRFCVSSSYSTQCSGPLREQRLCNNSAVCPVHG 410
QY      244  GNSWTWTEWSCSASCGRGWOKRSCRNTNPAPLNGAFCEGQNVHDTVSSLLV---SVDG 300
Db      411  AWDEWPWSCSCTCGGRFDRTICR--PPQFGNPGCEPKQTKFTCNALCFGRVDD 468

```

RESULT 4

TSHUP2
Thrombospondin 2 precursor - human
C:Species: Homo sapiens (man)
C:Date: 19-May-1995 #sequence_revision 03-Aug-1995 #text_change 13-Aug-1999
C:Accession: A47379; A42173
R:LaBell, T.L.; Byers, P.H.
Genomics 17, 225-229, 1993
A:Title: Sequence and characterization of the complete human thrombospondin 2 cDNA: pote
A:Reference number: A47379; MUID:94010892; PMID:8406456
A:Accession: A47379
A:Molecule type: mRNA
A:Residues: 1-1172 <LAB>
A:Cross-references: GB:L12350; NID:G307505; PIDN:AAA03703.1; PID:G307506
R:LaBell, T.L.; Milewicz, D.J.; Distche, C.M.; Byers, P.H.
Genomics 12, 421-429, 1992
A:Title: Thrombospondin II: partial cDNA sequence, chromosome location, and expression c
A:Reference number: A42173; MUID:92217961; PMID:1559694
A:Accession: A42173
A:Molecule type: mRNA

A;Residues: 560-1172 <LA2>
A;Cross-references: GB:M81339
A;Experimental source: fibroblast
A;Note: sequence extracted from NCBI backbone (NCBIN:95091, NCBIP:95096)
C;Genetics;
A;Gene: GDB:THBS2, TSP2
A;Cross-references: GDB:128789; OMIM:188061
A;Map position: 6q27-6q27
C;Complex: homotrimer, disulfide linked
C;Function:

C;Function:

A:Description: Participates in cell migration and adhesion, and in platelet aggregation
C:Superfamily: thrombospondin 1; EGF homology; thrombospondin type 1 repeat homology; von
C:Keywords: beta-hydroxyasparagine; calcium binding; cell adhesion; glycoprotein; trimer
F:1-18/Domain: signal sequence #status predicted <SIG>
F:19-1172/Product: thrombospondin 2 #status predicted <MAT>
F:319-377/Domain: von Willebrand factor type C repeat homology <VWC>
F:380-431/Domain: thrombospondin type 1 repeat homology <THR1>
F:436-492/Domain: thrombospondin type 1 repeat homology <THR2>
F:493-549/Domain: thrombospondin type 1 repeat homology <THR3>
F:553-588/Domain: EGF homology <EGF1>
F:652-691/Domain: EGF homology <EGF>
F:928-930/Region: cell attachment (R-G-D) motif
F:151,316,330,457,584,710,1069/Binding site: carbohydrate (Asn) (covalent) #status predicted
F:167-226/disulfide bonds: #status predicted
F:266,270/disulfide bonds: interchain #status predicted
F:612/Modified site: erythro-beta-hydroxyasparagine (Asn) #status predicted

Query Match	5.7%	Score 273.5;	DB 1;	Length 1172;
Best Local Similarity	28.9%	Pred. No. 1.3e-11;		

209	QY	QBARLATANYTCVAKNIIVARRRSASAA-VIYVYNGGWSTWEMSVCSASCGRGWQKRSR	267
403	Db	QGRSCDVNTNTCLGPSIQTRACSLSKCDTRIQDGGWSHWPSSSCSVTCGVGNITRIR	462
268	QY	SCTNPAPLNGAFCEGQVNHDRTVSSLLVSDGWSWPWSKWSACGLDCT---HWRSRECS	324
463	Db	LCNSPVPOMGGKNGKSGRETKACQAGAPCPIDGRWSPWPSACTVTCAGGIRETRVCN	522
325	QY	DPAPRNGEBCQGTDLTRNCTSLCVHSASGPEDVALYVGLIAVAVCLVLLLVLY	384
523	Db	SPBFPYGGKACGVQVBRQWCKRSC---PVDGCLSNPFCFGAQC-	564
385	QY	CRKKEGLSDVADSSILTSGFQPSVI--KSKADNPHLITQDPLSTTT-----TYQ	434
565	Db	-----SSFPDGS-WSCGFCPVGLNGHGHCEDECALVPDICFSTSKVPCVNTQP	615
435	QY	GSIC-----PRQDGFSP	446
616	Db	GFHCLPCPPRYRGNQP	631

RESULTS

thrombospondin 2 precursor - mouse
C:Species: Mus musculus (house mouse)
C:Date: 04-Mar-1993 #sequence_revision 18-Nov-1994 #text_change 20-Aug-1999
C:Accession: A42587; A39851
R:Laherty, C.D.; O'Rourke, K.; Wolf, F.W.; Katz, R.; Seldin, M.F.; Dixit, V.M.
J. Biol. Chem. 267, 3274-3281, 1992
A:Title: Characterization of mouse thrombospondin 2 sequence and expression during cell
A:Reference number: A42587; MUID:92147683; PMID:1371115
A:Accession: A42587
A:Status: preliminary; not compared with conceptual translation
A:Molecule type: nucleic acid
A:Residues: 1-1172 <LAH>
A:Cross-references: GB:I07803; GB:M87275; NID:g340421; PIDN:AAAS3064.1; PID:g567241
A:Note: sequence extracted from NCBI backbone (NCBIP:81502)
R:Bornstein, P.; O'Rourke, K.; Wikstrom, K.; Wolf, F.W.; Katz, R.; Li, P.; Dixit, V.M.
J. Biol. Chem. 266, 12821-12824, 1991
A:Title: A second, expressed thrombospondin gene (Thbs2) exists in the mouse genome.
A:Reference number: A39851; MUID:91302287; PMID:1712771
A:Accession: A39851

RESULT 9

T18856
angiogenesis inhibitor homolog - Caenorhabditis elegans
C:Species: Caenorhabditis elegans
C:Date: 15-Oct-1999 #sequence_revision 15-Oct-1999 #text_change 18-Feb-2000
C:Accession: T18856; T24653
R:McMurray, A.
submitted to the EMBL Data Library, July 1995
A:Reference number: Z19031
A:Accession: T18856
A:Status: preliminary; translated from GB/EMBL/DBJ
A:Molecule type: DNA
A:Residues: 1-1444 <W1>
A:Cross-references: EMBL:Z50004; PIDN:CAA90293.1; GSPDB:GN00028; CESP:C02B4.1
A:Experimental source: clone C02B4
R:McMurray, A.
submitted to the EMBL Data Library, July 1995
A:Reference number: Z19117
A:Accession: T24653
A:Status: preliminary; translated from GB/EMBL/DBJ
A:Molecule type: DNA
A:Residues: 1-1444 <W12>
A:Cross-references: EMBL:Z50006; PIDN:CAA90302.1; GSPDB:GN00028; CESP:C02B4.1
A:Experimental source: clone T07C5
C:Genetics:
A:Gene: CESP:C02B4.1
A:Map position: X
A:Introns: 25/3; 70/3; 96/3; 139/3; 187/1; 234/2; 282/3; 376/2; 422/2; 478/3; 509/3; 568

Query Match 5.5%; Score 263.5; DB 2; Length 1444;
Best Local Similarity 26.4%; Pred. No. 9.3e-11;
Matches 73; Conservative 26; Mismatches 91; Indels 87; Gaps 12;
QY 123 WCOCVWSSSGTTKSKAYIRIARLRKNFEQ-----EPLAKVSVLEQGVLPFCRPEGI 176
DB 1134 WSEWSSNSAC-----SCFSLTSTRRRFCQVVDPTVQGPCAGAILQ---IFCAPGSCS 1183
QY 177 PPAE-----VEW-----LRNEDLVPSLDPNVYITREHSLVVRQARLADTN 218
DB 1184 PSAGGWSLWSEWSSCKDCGDTGHQIRNRCSEF-----IPSNRGAYCSG 1228
QY 219 YT-----CYAKNIVARRRSASAAVIVYVNGSWTTEWSVCSACGRGWKRSRSTNPA 273
DB 1229 YSFQRPQVMDNVCSDEK-----VDGGWTDWTAWSECTDYCRNGHRSRTRFCANPK 1279
QY 274 PLMGGAFCGQN-----VHRTVSSLLVSDGWSWPSKWSACGLDC---THWR 319
DB 1280 PSQGGHACTGSDPELNPFCPPARCHLR-----DGGWSTWSDWTPCSASCFGVQTR 1330
QY 320 SRECSPPAPNGEECGQTDLTDRNCTSDLCVHSASG 356
DB 1331 DRSCSSPEPK-GGQCSGLAHQTSCLDLPACDHSOG 1366

RESULT 10

T00027
brain-specific angiogenesis inhibitor 2 - human
N:Alternate names: BAI2 protein
C:Species: Homo sapiens (man)
C:Date: 22-Jan-1999 #sequence_revision 22-Jan-1999 #text_change 21-Jul-2000
C:Accession: T00027
R:Shiratsuchi, T.; Nishimori, H.; Ichise, H.; Nakamura, Y.; Tokino, T.
Cytoget. Cell Genet. 79, 103-108, 1997
A:Title: Cloning and characterization of BAI2 and BAI3, novel genes homologous to brain
A:Reference number: Z14066; MUID:98194217; PMID:9533023
A:Accession: T00027
A:Status: translated from GB/EMBL/DBJ
A:Molecule type: mRNA
A:Residues: 1-1572 <SHI>
A:Cross-references: EMBL:AB005298; NID:g3021698; PIDN:BAA25362.1; PID:g3021699
A:Experimental source: brain
C:Genetics:

A:Gene: GDB:BAI2
A:Cross-references: GDB:9838089; OMIM:602683
A:Map position: lp35-lp35

Query Match 5.4%; Score 258; DB 2; Length 1572;
Best Local Similarity 18.8%; Pred. No. 2.6e-10;
Matches 172; Conservative 110; Mismatches 312; Indels 322; Gaps 37;

QY 173 PEGIPPAEVEWURNEDLVDPSPDNVY-----ITREHSLVVRQARL 213
DB 271 PEEEPKVKTPRSAD-----EPGLYMAQTGDPAAEWSWPSVCSLTCCGGLQVR-TKS 323
QY 214 ADTANYTCVAKNIVARRRSASAAVIVYVNGSWTTEWSVCSACGRGWKRSRSTNPA 273
DB 324 CVSSPYGLCSPLRETRPCNNSATCPVHGWEWGSNSLCSRSCGRGSRMRITCV--P 381
QY 274 PLNGGAFCGQNVDRTVSSLLVSDGWSWPSKWSACGLDC---THWRSRCSDDPAPR- 329
DB 382 PQHGGKACGPELQTKLCMAACPVEGQWLEWPGWPCSTSCANGTQQRSKCSVAGPAW 441
QY 330 -----NGEECQ 336
DB 442 ATCTGALTDRCSNLCEPATDSKMGPMNWSLCSKTCDTGMQRFRMCOATGTQYPCE 501
QY 337 GTDLDRNCTSDLC--VHSASGPELVAL----- 362
DB 502 GTGEEVKPCSEKRCFAFHEMCRDEYVMLMTWKAAGEIYINKPCPNASGSASRRCLLSA 561
QY 363 ----YVGLJAVACL---VLLLVLLVYCRKKEGLSDVADSSILTSQFPVSIKPSKA 415
DB 562 QGVAYWGLPSFARCISHREYRYLYLSREHLAKQRMAGEGMSQVVR-SLQELLARITY 620
QY 416 DNPILLTIQPDLSITTY--QGSCLCPRODQSPSPKFLT-----NGHLLPLGG 461
DB 621 SGDLFSVILRNVDITFKRATVPSADVDQRFQVVSFMDAENKEKMDDAQVSP--G 678
QY 462 GRHTLHSSPTSEAEFV-----SRLSNTYFRSLPRG-----TSNMTYGTEN 504
DB 679 SVHLR-----VVEFHLVGDALAKAFQSSLLVTDNLVISIQREPVSASSDITPMRG 732
QY 505 FLG-----GRLMIPNTGISLLIP-----PDALPRGK----- 530
DB 733 RRGMDKDWHRSEDRFLPKPEVLSLSPGKPATSGAAGSPGRGPGGTVPDPGSHQRL 792
QY 531 -----IYE-IYLTILHKPEVDRLPLACQTLSPVSCGPGVLLTRPVIL 574
DB 793 PADPDESSYFVIGAVLYRTLGLILPPP---RPPLAVTSRYMT--VTVRPPTQPPAEPLIT 847
QY 575 A-----MDHCGEPPDSWSLRKQSCGSEWQDVLHGEAPSHLYYCOLEASACVFT 629
DB 848 VELSYIITGTTDPHCASWDYS--RADASSGDWDTEHCOTLETOAHT-RCQ-----C 896
QY 630 EQLGRFALVGE-----ALSVAARLKLALLFAPVACTSLEYNIRVYCLHDTDALKEVVQ 684
DB 897 QHLSLFAVLAQPPKDLTLELAGSPSVPLVIGCAVSCWALLTLAIYA-----AFWRFIK 950
QY 685 LEKQLGGLIQEPRVLFKDSYHNRLSIHDVPSSLWKSLLYSYOEIPFYHLWNGTQRY 744
DB 951 SERSI-----ILNFCLSI--LASNI---LILVQSRVLSKGVCTMTAAF 990
QY 745 LHCTFTLERSVPSSTDLACKLWV-----WQVEGDGQS 776
DB 991 LHFFF-----LSSFCWLTWQSYLAVIGEMRTRLYRKRFLCLGWGLPALVVA 1039
QY 777 FSINFNITKDTRFAPALLALESEAG-VPALVGPQA-----FKTFFLIROKI-----ISSL 824
DB 1040 VSVGFTRTKGYGFSYCWLSLEGLGLLYFVGPAVIVLVNMLGIIIVFNKLMARDGISDK 1099
QY 825 DPPCRGAD---WRTL 837
DB 1100 SKQKAGASRCPWASL 1115

F:926-928/Region: cell attachment (R-G-D) motif
F:171-232/Disulfide bonds: #status predicted
F:248,360,708,1067/Binding site: carbohydrate (Asn) (covalent) #status predicted
F:270,274/Disulfide bonds: interchain #status predicted
F:610/Modified site: erythro-beta-hydroxyasparagine (Asn) #status predicted
F:1051/Binding site: carbohydrate (Asn) (covalent) #status absent

Query Match 5.1%; Score 246; DB 1; Length 1170;
Best Local Similarity 30.3%; Pred. No. 1.3e-09;
Matches 47; Conservative 24; Mismatches 66; Indels 18; Gaps 3;

QY 207 VVROARLADTANYTCVAKNIVAR-----RRSASAAVIVVVGWSTWTWVSVCASG 258
Db 399 IQGRGSCDLSNNRCEGSSVQTRTCHIQCDDRKFKQ-----DGGNSHSPWSSCSVTC 451
QY 259 GRGWQXRSRCTNPAPINGAFCEGQNVHRTVSSLLVSDGWSWSPWKSACGLDC--- 315
Db 452 GDGVITRILCNSPSPQMGKPCGEARETKACKKDACPDGNGWSPWDICSVTCGGG 511
QY 316 THWRSRECDPAPRNGECQGTDLTRNCTSDLC 350
Db 512 VOKRSRLCNPPTQFGGKDCVGDVTENQCNKQDC 546

RESULT 13
A39804
thrombospondin precursor - chicken
C:Species: Gallus gallus (chicken)
C:Date: 10-Sep-1999 #sequence_revision 10-Sep-1999 #text_change 10-Sep-1999
C:Accession: A39804
R:Lawler, J.; Duguet, M.; Ferro, P.
J. Biol. Chem. 266, 8039-8043, 1991
A:Title: Cloning and sequencing of chicken thrombospondin.
A:Reference number: A39804; MUID:91217026; PMID:2022631
A:Accession: A39804
A:Status: preliminary
A:Molecule type: mRNA
A:Residues: 1-1178 <LAW>
A:Cross-references: GB:M60853; NID:g212763; PID:AAA51437.1; PID:g212764
C:Superfamily: thrombospondin 1; EGF homology; thrombospondin type 1 repeat homology; v
F:325-383/Domain: von Willebrand factor type C repeat homology <VMC>
F:386-437/Domain: thrombospondin type 1 repeat homology <THR1>
F:442-498/Domain: thrombospondin type 1 repeat homology <THR2>
F:499-555/Domain: thrombospondin type 1 repeat homology <THR3>
F:658-697/Domain: EGF homology <EGF>

Query Match 5.1%; Score 243.5; DB 1; Length 1178;
Best Local Similarity 34.4%; Pred. No. 2.1e-09;
Matches 55; Conservative 15; Mismatches 75; Indels 15; Gaps 4;

QY 210 QARLADTANYTCVAKNIVARRRS-ASAAVIVVYNGHSTWTSVCSASCGRQKRSRS 268
Db 410 RGRSCDVTRSACTGPHIQTRMCSFKKCDHRIRQDGGWSHSPWSSCSVTCGVGNITRIL 469
QY 269 CTNPAPINGAFCEGQNVHRTVSSLLVSDGWSWSPWKSACGLDC---THWRSRECD 325
Db 470 CNGPIPMGKGNVGNRETEKCEKAPCPVNGQWGPWSPWSACTVTCGGIRSRSLCNS 529
QY 326 PAPRNGEECQGTDLDT-----RNCTSDLCVHSAAGP 357
Db 530 PEQYGGKPCVG---DTKQHDMCNKRDCPDGCLSNFCFP 566

RESULT 14
T14160
transmembrane receptor protein Robo1 - rat
C:Species: Rattus norvegicus (Norway rat)
C:Date: 20-Sep-1999 #sequence_revision 20-Sep-1999 #text_change 20-Sep-1999
C:Accession: T14160
R:Kidd, T.; Brose, K.; Mitchell, K.J.; Fetter, R.D.; Tessier-Lavigne, M.; Goodman, C.S.;
Cell 92, 205-215, 1998
A:Title: Roundabout controls axon crossing of the CNS midline and defines a novel subfam
A:Reference number: Z17897; MUID:98117249; PMID:9458045

A:Accession: T14160
A:Status: preliminary; translated from GB/EMBL/DBJ
A:Molecule type: mRNA
A:Residues: 1-1651 <KID>
A:Cross-references: EMBL:AF041082; NID:g2811215; PID:g2811216; PID:AA39960.1
C:Function:
A:Description: appears to function as the gatekeeper controlling midline crossing
C:Keywords: transmembrane protein

Query Match 4.6%; Score 220.5; DB 2; Length 1651;
Best Local Similarity 20.1%; Pred. No. 1.6e-07;
Matches 205; Conservative 132; Mismatches 377; Indels 305; Gaps 48;

QY 41 DLLPHFLVEPDVIVKKNKPVLLCKAV--PATQFFKCKNGEWMROVDHVIERS----- 92
Db 65 DPFPRIVEHPSDLIVSKGEPATLNCKAQRPTTIEWYKGERV-ETDKDDPRSHMLLP 123
QY 93 -----TDGSSGEPTMEVRINVSROQVEKVFGLLEYWCQCVAMSSGGTTKSKAYIR 143
Db 124 SGLSFLFRLIVHGRKSRPDGVYI-----CVARNYLGEAVSHNASLE 164
QY 144 IARLKNFQEBPLAKEVSLQGIIVLPCRPEGIPPAEVEWLNEDLVDPISLPNVYIT-R 202
Db 165 VAILRDDFRQNPSDVMVAVGEPAVMCEQPGRGHPPTTISKKDGGPLD---DKDKRITIR 221
QY 203 EHSLVVRQARLADTANYTCVAKNIVARRRSASAAVIVVYNGW-----STWTSV-- 253
Db 222 GGLMITYTRKSDAGKVCVGNMGERESKVDVTLERPSFVKRPSNLAVTVDDSAEF 281
QY 254 -CSASGR---GWQK-----RSRCTNPAPINGAFCEGQNVHRT 290
Db 282 KCEARGDPVPTFGWRKDDGELPKSRYEIRDHTLKIRKVT--AGDMGSYTCVAENMVGA 339
QY 291 VSLLSVVDGS-----WSPWKSACGLDCTHWRSECSAP-----RNGGEC----- 335
Db 340 EASATLTQEPHFVVKPRDQVVALGRVT--FQCEATGNPQPAIFWRREGSQMLLSYQP 398
QY 336 -QGTDLDTNCTSDLCVHSAAGPVDVAVGLIAVAVCLVLLVLLVLYCRKKEGLDSD 394
Db 399 POSRSFVSQGLDVTNVQ-RSDVGY-----ICQTL-----N 432
QY 395 VADSSILTSQFQPVSIKPSKADNPHLLTIQPDLSITTTTQGS-----CPRQDGSFPK- 448
Db 433 VA-GSIITKAY--LEVTDLADRPPIRQGVNQTAVDGTTLTSCVATGSPVPTILWR 489
QY 449 -----QLTNCHL-----LSPLG-GRHTLHSSPTSEAEFVSRLSTQNYFRS 490
Db 490 KQGLVSTQDSRIKQLESGVLQIRYAKLGDGTGRYTCTASTPSGEA----- 534
QY 491 LPRGTSNMTYGTENFLGRLMIENGTISLLIP-----PDAIPRGKIYEIYTLHKPEDVRL 546
Db 535 -----TWSAY-----IEVQEGVPVQPPRPTDNLIPSAP-----SKPEVTDV 572
QY 547 PLACQTLTASPIVSCGPPGVLLTRPVLAMDHGCEPSPDSWLSRLKKQSCGEGWEQDLVH 606
Db 573 SKNTVTLLWQPNL--SGATPTSIIIEAFSHA---SGSSW-----QTVAENVKTETFA 620
QY 607 LGEEAPSHLYYQLEASACVYFTE-----QIGRFALVG 639
Db 621 IKGLKPNAIYFLVRAANAYGISDPVKTQDVPPPTQGVHDHKKVQVRELG----- 674
QY 640 EALSVAARKLKLFPAPVACTSLEYNIRVYCLHDTHALKEVVVQLEKQLGGOLIOBPRV 699
Db 675 -----NVVLHNLNPTILSSSVYEV-----HWITVDQOSQYIQ-----GYKILYRPS- 714
QY 700 LHFKDSYHNLRSLIHDVPSSLWKSULL-----VSYQ--BIPFYHIWNGQRYLHCTFTL 751
Db 715 ---GASGESEWLVEVTRPTKNSVVIPLDKRGVNYEIKARPFENFQGDADSEIKFAKTL 771
QY 752 ERVSPSTSLACKLWVQWEGDQGSINFIN-ITKOTRFAEL-----LALSEACGVA 803
Db 772 EE-RPSAPPSS--VTVSKNDGNGTALVTWQPPEDTQNGVVOEYKVMCLGNETRYHINK 828

QY 804 LVGPSAFK--IPFLIRQKIISLLDPPCRRGA--DWRTLAQKLIHLDH----- 846
Db 829 TVDGSFTSVVIFPLV-PCIRYSVEVAASTGAGPGVKSEPOFIQLDHSNPNVSPEDQVSLA 887
QY 847 --LSFFASKSPPTA-----MILNLWEARHFPNGNUSQLAAAVAGLGQPDAGLFT 893
Db 888 QQISDVVVKQPAFTAGIAGAACWIIILMVFSIWLVRHKKRN--GLSSTVAGIRKVPSTFT 944

RESULT 15
T30805
dutt1 protein - mouse
N;Alternate names: transmembrane receptor protein Robol homolog
C;Species: Mus musculus (house mouse)
C;Date: 22-Oct-1999 #sequence_revision 22-Oct-1999 #text_change 22-Oct-1999
C;Accession: T30805
R;Wu, M.C.; Lowe, N.; Fordham, R.; Rabbitts, P.
submitted to the EMBL Data library, July 1998
A;Description: The mouse homologue of human DUTTI/H-robol gene: protein sequence and chr
A;Reference number: Z20879
A;Accession: T30805
A;Status: preliminary; translated from GB/EMBL/DBJ
A;Molecule type: mRNA
A;Residues: 1-1612 <NUM>
A;Cross-references: EMBL:Y17793; NID:el329712; PID:el329713; PIDN:CAA76850.1
C;Genetics:
A;Gene: dutt1
A;Map position: 16

Query Match 4.5%; Score 217.5; DB 2; Length 1612;
Best Local Similarity 19.9%; Pred. No. 2.6e-07;
Matches 211; Conservative 133; Mismatches 359; Indels 355; Gaps 52;

QY 41 DLLPHFLVPEPVYVKNKPVLLVCKAV--PATQIFPKMGWVRQVDHVIERS----- 92
Db 26 DPPPRIVEHPSDLIVSGEPATLNCKAAGRPPTTIEWYKGERV-ETDKDDPRSHRMLLP 84
QY 93 -----TDGSSGEPTMEVRINVRQVKEVFGLEEYWCQCVWSSSGTTKSKAYIR 143
Db 85 SGSLFFLRIVHGKSRPDEGVYI-----CVARNYLGEAVSHNASLE 125
QY 144 IARLRNFOEPLAKEVSLQGVILPCRPPEGIPPAEVEWLRNEDLVPSLDENVYIT-R 202
Db 126 VAILRDDFRQNSDVNMVAVGEPAVMECPQPRGHEPTISWKDGSPLD--DKDERITIR 182
QY 203 EHSLVVRQARLADTANYTCVAKNIVARRBSASAAVIVYNGW-----STWTEWSV-- 253
Db 183 GGLMLITYTRKSDAGKYVCVTNMGERESEVAELTVLERPSEVKRPSNLAVTVDDSAEF 242
QY 254 -CSASGGRG-----WQK-----RSRSCTNPAPLNGGAPCEGQNVH 287
Db 243 KCEA---RGDPVPTVRWKDDGELPKSRYEIRDDHTLKIRKVT--AGDMGSYTCVAENNV 297
QY 288 DRTVSSLLVSDGS-----WSPWSKWSAGLDCTHWSRECSDPAP-----RNGGEC--- 335
Db 298 GKAEASATLTVOEPHFVVKPRDQVVALGRVT-FOCEATGNPQPAIFWRREGSQNLLFS 356
QY 336 ----QGTDLDRNCTSLDLCVHSGASGEDVALYVGLVAVCLVLLVLLVLYCKKKEGL 391
Db 357 YQPPQSSSRFSVQSGDLTITNVQ-RSDVGYI-----ICQTL----- 392
QY 392 DSDVADSSILTSGGQFVSIKPSKADNPHELLTIQPDLSITTTTYQGSIL---CPRQDGSPKPF 448
Db 393 --NVA-GSIITKAY--LEVTQVADRPPVIRQGPVQVAVDGTILSCVATGSPAPTI 447
QY 449 -----QLTNGHL---LSPLGG-GRHTLHSSPTSEAEFEVSRSLSTQNY 487
Db 448 LMRKQGVINVSQDSRIKQLESGLQIRYAKLGDTGRYCTCTASTPSGEA----- 495
QY 488 FRSLPRGTSNMTYGTGFNFGGRIMIPNTGISLLIP-----PDAIPRGKIYEIYLHLKPE- 542
Db 496 -----TWSAY-----LEVOEGFVQVPPRPTDNLIPSAP-----SKPEV 530

QY 543 -DV---RLPLAGCOTLLSPIVSCPPGVLLTRPVILAMDHCGEPSPDWSLRLLKKQCEG 598
Db 531 TDVSKNTVTILSWQPNLNS-----CATPTSYIIIEAFSHA---SGSSW-----QTAAE 573
QY 599 SWEQDVILHGEAPSHLYYCQLERASACVVTE-----Q 631
Db 574 NVKTETFAIKGLKPNALYFLVRAANAYISDPESQISDPVKTDQVPTSQGVHKKVQRE 633
QY 632 LGRFALVGEALSVAALKRLKLLFAPVACTSLEYNIERYVCLHDTHDALKEVVQLEKQLGG 691
Db 634 LG-----NVVLHLHNPILSSSSEV-----HWTVDQOSQYIQ-----GY 668
QY 692 QLIQEPRLVLPKDSYHNLRSLIHDPSSLWKSLL-----VSYQ--EIPFYHIWNGTOR 743
Db 669 KILYRPS-----GASHGESEWLVEVRTPTKNSVVIPDLRKGVNVEIKARPPFFNEFOGADS 724
QY 744 YLHCTFTLERVSPSTSLACKLWVWQVGGQSFNEN-ITKDTREAFEL-----LAL 795
Db 725 EIKFAKTLEB-AFSAPRS--VTVSKNDGNGTALVTWQPPEDTQNGMVQEVYKWCICGN 781
QY 796 ESEAGVPALVGPFAKIPFLIRQKIISLDDPPCR-----RGADWRTLAQKLHDS 845
Db 782 ETKYHINKTVDGSTFSV-----VIPSLVPGIRYSVEVAASTGAGPGVKSEPOFIQLDS 834
QY 846 H-----LSFFASKSPPTA-----MILNLWEARHFP--NGNLSQLA 878
Db 835 HGNVSPEDQVSLAQOISDVVRQPAFIAGIAGAACWIIILMVFSIWLVRHKKRNLGTSTYA 894
QY 879 A-----AVAGLGQPDAGLFTVSE 896
Db 895 GIRKVPSTFTPTVTYQRGGEAVSSGGRP--GLLNISE 930

Search completed: October 4, 2004, 18:32:42
Job time : 56 secs

GenCore version 5.1.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - protein search, using sw model

Run on: October 4, 2004, 18:11:25 ; Search time 27 Seconds
(without alignments)

1733.743 Million cell updates/sec

Title: US-09-970-944-2

Perfect score: 4787

Sequence: 1 MAVRGLPALLGIVLAARWL.....AVAGLQPDAGLFTVSEAEK 899

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 141681 seqs, 52070155 residues

Total number of hits satisfying chosen parameters: 141681

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : SwissProt_42.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	280	5.8	1584	1 BAI1 HUMAN	O14514 homo sapien
2	273.5	5.7	1172	1 TSP2 HUMAN	P35442 homo sapien
3	273.5	5.7	1172	1 TSP2 MOUSE	Q03350 mus musculus
4	272.5	5.7	1077	1 SM5A MOUSE	Q62217 mus musculus
5	270.5	5.7	1074	1 SM5A HUMAN	Q13591 homo sapien
6	269	5.6	1170	1 TSP2 BOVIN	Q95116 bos taurus
7	267.5	5.6	1522	1 BAI3 HUMAN	O60242 homo sapien
8	266.5	5.6	1173	1 TSP1 XENLA	P35448 xenopus lae
9	258.5	5.4	1093	1 SM5B MOUSE	O60519 mus musculus
10	258	5.4	1572	1 BAI2 HUMAN	O60241 homo sapien
11	252	5.3	1093	1 SM5B HUMAN	Q9p283 homo sapien
12	248	5.2	1170	1 TSP1 MOUSE	P35441 mus musculus
13	246	5.1	1170	1 TSP1 HUMAN	P07996 homo sapien
14	243.5	5.1	1178	1 TSP2 CHICK	P35440 gallus gall
15	243	5.1	1170	1 TSP1 BOVIN	Q28178 bos taurus
16	208.5	4.4	469	1 PROP_HUMAN	P27918 homo sapien
17	205.5	4.3	470	1 PROP_CAVPO	Q84181 cavia porce
18	202.5	4.2	437	1 PROP_MOUSE	P11680 mus musculus
19	186	3.9	1266	1 NGCA CHICK	Q03696 gallus gall
20	181.5	3.8	867	1 SSPO BOVIN	P98167 bos taurus
21	178.5	3.7	1223	1 AT14 HUMAN	O8wxs8 homo sapien
22	175.5	3.7	1736	1 ZOI1 HUMAN	Q07157 homo sapien
23	172	3.6	630	1 AT5A_RAT	Q9esp7 rattus norv
24	170.5	3.6	1745	1 ZOI1 MOUSE	P39447 mus musculus
25	169	3.5	934	1 CO6 HUMAN	P13671 homo sapien
26	167.5	3.5	905	1 AT5B MOUSE	P57110 mus musculus
27	166	3.5	837	1 AT5A HUMAN	O75173 homo sapien
28	162.5	3.4	587	1 CO8B_ONCMY	Q90x85 oncorhynch
29	160.5	3.4	562	1 AT15 MOUSE	P59384 mus musculus
30	160.5	3.4	967	1 AT10 HUMAN	Q9uh18 homo sapien
31	160.5	3.4	1077	1 AT10 HUMAN	Q9h324 homo sapien
32	159.5	3.3	1224	1 AT16 HUMAN	O8te57 homo sapien
33	156.5	3.3	860	1 AT56_HUMAN	Q9ukp5 homo sapien

RESULT 1

ID	BAI1_HUMAN	STANDARD;	PRT;	1584 AA.
AC	O14514;			
DT	16-OCT-2001 (Rel. 40, Created)			
DT	16-OCT-2001 (Rel. 40, Last sequence update)			
DT	28-FEB-2003 (Rel. 41, Last annotation update)			
DE	Brain-specific angiogenesis inhibitor 1 precursor.			
GN	BAI1.			
OS	Homo sapiens (Human).			
OC	Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;			
OC	Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.			
OX	NCBI_TaxID=9606;			
RN	[1]			
RP	SEQUENCE FROM N.A.			
RC	TISSUE=Fetal brain;			
RX	MEDLINE=98054121; PubMed=9393972;			
RA	Nishimori H., Shiratsuchi T., Urano T., Kimura Y., Kiyono K.,			
RA	Tatsumi K., Yoshida S., Ono M., Kuwano M., Nakamura Y., Tokino T.;			
RT	"A novel brain-specific p53-target gene, BAI1, containing			
RT	thrombospondin type 1 repeats inhibits experimental angiogenesis.";			
RL	Oncogene 15:2145-2150(1997).			
RN	[2]			
RP	INTERACTION WITH BAPI.			
RX	MEDLINE=98321173; PubMed=9647739;			
RA	Shiratsuchi T., Futamura M., Oda K., Nishimori H., Nakamura Y.,			
RA	Tokino T.;			
RT	"Cloning and characterization of BAI-associated protein 1: a PDZ			
RT	domain-containing protein that interacts with BAI1.";			
RL	Biochem. Biophys. Res. Commun. 247:597-604(1998).			
CC	-!- FUNCTION: LIKELY TO BE A POTENT INHIBITOR OF ANGIOGENESIS IN			
CC	BRAIN AND MAY PLAY A SIGNIFICANT ROLE AS A MEDIATOR OF THE P53			
CC	SIGNAL IN SUPPRESSION OF GLOBLASTOMA. MAY FUNCTION IN CELL			
CC	ADHESION AND SIGNAL TRANSDUCTION IN THE BRAIN.			
CC	-!- SUBUNIT: INTERACTS WITH BAPI.			
CC	-!- SUBCELLULAR LOCATION: INTEGRAL MEMBRANE PROTEIN. LIKELY TO BE			
CC	CONCENTRATED AT CELL-CELL ADHESION SITES.			
CC	-!- TISSUE SPECIFICITY: SPECIFICALLY EXPRESSED IN BRAIN. REDUCED OR NO			
CC	EXPRESSION IS OBSERVED IN SOME GLIOBLASTOMA CELL LINES AND CANCER			
CC	TISSUES.			
CC	-!- INDUCTION: By p53.			
CC	-!- DOMAIN: THE TSP1 REPEATS INHIBIT IN VIVO ANGIOGENESIS IN RAT			
CC	CORNEA INDUCED BY BFGF.			
CC	-!- SIMILARITY: Belongs to family 2 of G-protein coupled receptors.			
CC	-!- SIMILARITY: Contains 5 TSP type-1 domains.			
CC	-!- SIMILARITY: Contains 1 GPS domain.			
CC	-----			
CC	This SWISS-PROT entry is copyright. It is produced through a collaboration			
CC	between the Swiss Institute of Bioinformatics and the EMBL outstation -			
CC	the European Bioinformatics Institute. There are no restrictions on its			
CC	use by non-profit institutions as long as its content is in no way			
CC	modified and this statement is not removed. Usage by and for commercial			
CC	entities requires a license agreement (see http://www.isb-sib.ch/announce/)			
CC	or send an email to license@isb-sib.ch .			
CC	-----			
CC	EMBL; AB005297; BAA23647.1; -.			

Q8te58 homo sapien
Q9ukp4 homo sapien
P58397 homo sapien
Q8c9w3 m adamcs-2
O15072 homo sapien
Q9wug1 rattus norv
P97857 mus musculus
Q9up79 homo sapien
Q8te56 homo sapien
Q9una0 homo sapien
Q98902 fugu rubrip
P55314 rattus norv

ALIGNMENTS

DR PIR; T00026; T00026.
 DR Genew; HGNC:943; BAIL.
 DR MTM; 602682; -.
 DR GO; GO:0005897; C: integral to plasma membrane; TAS.
 DR GO; GO:0005911; C: intercellular junction; TAS.
 DR GO; GO:0005911; F: protein binding; TAS.
 DR GO; GO:0007409; P: axonogenesis; TAS.
 DR GO; GO:0007422; P: cell adhesion; TAS.
 DR GO; GO:0008285; P: negative regulation of cell proliferation; TAS.
 DR GO; GO:0008285; P: peripheral nervous system development; TAS.
 DR GO; GO:0007422; P: signal transduction; TAS.
 DR GO; GO:0007422; P: signal transduction; TAS.
 DR InterPro; IPR000832; GPCR secretin.
 DR InterPro; IPR001879; hormu receptor.
 DR InterPro; IPR002023; PKD cys_rich.
 DR InterPro; IPR000884; TSP1.
 DR Pfam; PF00002; 7tm_2; 1.
 DR Pfam; PF01825; GPS; 1.
 DR Pfam; PF02793; HRM; 1.
 DR Pfam; PF00090; tsp_1; 5.
 DR SMART; SM00303; GPS; 1.
 DR SMART; SM00008; Hormr; 1.
 DR SMART; SM00209; TSP1; 5.
 DR PROSITE; PS00221; GPS; 1.
 DR PROSITE; PS00649; G PROTEIN RECP F2_1; FALSE NEG.
 DR PROSITE; PS00650; G PROTEIN RECP F2_2; FALSE NEG.
 DR PROSITE; PS00227; G PROTEIN RECP F2_3; 1.
 DR PROSITE; PS00261; G PROTEIN RECP F2_4; 1.
 DR PROSITE; PS00092; TSP1; 5.
 KW G-protein coupled receptor; Transmembrane; Glycoprotein; Signal;
 KW Repeat; Cell adhesion.
 FT SIGNAL 1 30
 FT CHAIN 31 1584
 FT DOMAIN 31 948
 FT TRANSMEM 949 969
 FT DOMAIN 970 980
 FT TRANSMEM 981 1001
 FT DOMAIN 1002 1008
 FT TRANSMEM 1009 1029
 FT DOMAIN 1030 1052
 FT TRANSMEM 1053 1073
 FT DOMAIN 1074 1093
 FT TRANSMEM 1094 1114
 FT DOMAIN 1115 1136
 FT TRANSMEM 1137 1157
 FT DOMAIN 1158 1166
 FT TRANSMEM 1167 1187
 FT DOMAIN 1188 1584
 FT TRANSMEM 261 315
 FT DOMAIN 354 407
 FT TRANSMEM 409 462
 FT DOMAIN 467 520
 FT TRANSMEM 522 575
 FT DOMAIN 581 938
 FT TRANSMEM 1411 1422
 FT DOMAIN 1425 1430
 FT TRANSMEM 1430 1430
 FT SITE 231 233
 FT TRANSMEM 1365 1584
 FT DOMAIN 1581 1584
 FT CARBOHYD 64 64
 FT CARBOHYD 401 401
 FT CARBOHYD 607 607
 FT CARBOHYD 692 692
 FT CARBOHYD 844 844
 FT CARBOHYD 877 877
 FT CARBOHYD 881 881
 FT SEQUENCE 1584 AA; 173531 MW; DEAF28C77874513 CRC64;
 Query Match 5.8%; Score 280; DB 1; Length 1584;
 Best Local Similarity 31.8%; Pred. No. 3.8e-13;
 Matches 74; Conservative 38; Mismatches 93; Indels 28; Gaps 10;
 QY 124 CQCWSSSGTYSQAYTRIALRKNPEQEPFLAKEVLSQGVLPFCRPFEGIPPAEVEW 183
 DR PIR; A47379; TSHUP2.

Db 309 CNREACGPAGRTSSRSQSLRSTARR---REELGDEL---QQGFPA-PQTGDPAAE-EW 360
 QY 184 LRNEDLVDPDLDPNVITRHSLSVVRQARLADTANTYTCVAKNIVARRRSASAAVIVYVNG 243
 Db 361 --SPWSVCSSTCGEGWQTR-----TRFCVSSYSTQCSGLPRLQCNNSAVCPVHG 410
 QY 244 GWSTWTWMSVCSASCCGRGWQKRSCTNPAPLNGAFCEGQNVHDRVSSLLV---SVDG 300
 Db 411 AWDEWSPWLSCTSCGGRFRDTRTCR--PPQFGNFCGEPKQTKFCNTALCPGRAVDG 468
 QY 301 SWSFWSKWSACGLDCT---HWRRECSDPAPRNGRERCCQGTDLDRNCTSDLC 350
 Db 469 NWNESSWSACSASCGRQORTRECNGPS--YGGACQGHWTETDCFLQOC 519
 RESULT 2
 TSP2 HUMAN
 ID TSP2 HUMAN STANDARD; PRT; 1172 AA.
 AC P35442;
 DT 01-JUN-1994 (Rel. 29, Created)
 DT 01-JUN-1994 (Rel. 29, Last sequence update)
 DT 28-FEB-2003 (Rel. 41, Last annotation update)
 DE Thrombospondin 2 precursor.
 GN THBS2 OR TSP2.
 OS Homo sapiens (Human).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
 OC NCBI_TaxID=9606;
 RN [1]
 RP SEQUENCE FROM N.A.
 RX MEDLINE=94010892; PubMed=8406456;
 RA Labell T.L., Byers P.H.;
 RT "Sequence and characterization of the complete human thrombospondin 2
 cDNA: potential regulatory role for the 3' untranslated region.";
 RL Genomics 17:225-229(1993).
 RN [2]
 RP SEQUENCE OF 560-1172 FROM N.A.
 RX TISSUE=Fibroblast;
 RA Labell T.L., McGookey Milewicz D.J., Distche C.M., Byers P.H.;
 RT "Thrombospondin II: partial cDNA sequence, chromosome location, and
 expression of a second member of the thrombospondin gene family in
 humans.";
 RL Genomics 12:421-429(1992).
 RN [3]
 RP THROMBOSPONDIN REPEATS DISULFIDE BONDS.
 RX MEDLINE=21588233; PubMed=11590138;
 RA Misenerheimer T.M., Hahr A.J., Harms A.C., Annis D.S., Mosher D.F.;
 RT "Disulfide connectivity of recombinant C-terminal region of human
 thrombospondin 2.";
 RL J. Biol. Chem. 276:45882-45887(2001).
 CC -!- FUNCTION: Adhesive glycoprotein that mediates cell-to-cell and
 cell-to-matrix interactions. Can bind to fibrinogen, fibronectin,
 laminin and type V collagen.
 CC -!- SUBUNIT: Homotrimer; disulfide-linked.
 CC -!- SIMILARITY: Belongs to the thrombospondin family.
 CC -!- SIMILARITY: Contains 1 WWFC domain.
 CC -!- SIMILARITY: Contains 3 EGF-like domains.
 CC -!- SIMILARITY: Contains 3 TSP type-1 domains.
 CC -!- SIMILARITY: Contains 7 TSP type-3 domains.
 CC -!- SIMILARITY: Contains 1 TSP N-terminal (TSPN) domain.
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 between the Swiss Institute of Bioinformatics and the EMBL Outstation -
 the European Bioinformatics Institute. There are no restrictions on its
 use by non-profit institutions as long as its content is in no way
 modified and this statement is not removed. Usage by and for commercial
 entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 or send an email to license@isb-sib.ch).

 EMBL; L12350; AAA03703.1; -.
 DR EMBL; M81339; -; NOT_ANNOTATED_CDS.
 DR PIR; A47379; TSHUP2.

```

DR HSP: P00740; 1EDM.
DR Genew; HGNC:11786; THBS2.
DR GO: 0008201; F: heparin binding; TAS.
DR MIM: 188061; -.
DR InterPro; IPR001881; EGF Ca.
DR InterPro; IPR006209; EGF like.
DR InterPro; IPR006210; IEGF.
DR InterPro; IPR000884; TSP1.
DR InterPro; IPR008085; TSP 1.
DR InterPro; IPR003367; tsp 3.
DR InterPro; IPR008089; TSPC.
DR InterPro; IPR003129; TSPN.
DR InterPro; IPR001007; WFC_C.
DR Pfam; PF00008; EGF; 2.
DR Pfam; PF00090; tsp_1; 3.
DR Pfam; PF02412; tsp_3; 13.
DR Pfam; PF05735; TSPN; 1.
DR Pfam; PF02210; TSPN; 1.
DR Pfam; PF00093; wvc; 1.
DR PRINTS; PRO1705; TSP1REPEAT.
DR SMART; SM00181; EGF; 3.
DR SMART; SM00209; TSP1; 3.
DR SMART; SM00210; TSPN; 1.
DR SMART; SM00214; WVC; 1.
DR PROSITE; PS00022; EGF 1; FALSE_NEG.
DR PROSITE; PS01186; EGF_2; 1.
DR PROSITE; PS00026; EGF_3; 2.
DR PROSITE; PS00092; TSP1; 3.
DR PROSITE; PS01208; WFC 1; 1.
DR PROSITE; PS01084; WFC_2; 1.
KW Glycoprotein; Cell adhesion; Calcium-binding; Heparin-binding; Repeat;
KW EGF-like domain; Signal.
FT SIGNAL 1 18
FT CHAIN 19 1172
FT DOMAIN 19 215
FT DOMAIN 19 232
FT DOMAIN 318 375
FT DOMAIN 381 431
FT DOMAIN 437 492
FT DOMAIN 494 549
FT DOMAIN 549 589
FT DOMAIN 590 647
FT DOMAIN 648 692
FT DOMAIN 725 760
FT DOMAIN 761 783
FT DOMAIN 784 819
FT DOMAIN 820 842
FT DOMAIN 843 880
FT DOMAIN 881 916
FT DOMAIN 917 952
FT DOMAIN 953 1172
FT SITE 928 930
FT DISULFID 266 266
FT DISULFID 270 270
FT DISULFID 393 425
FT DISULFID 397 430
FT DISULFID 408 415
FT DISULFID 449 486
FT DISULFID 453 491
FT DISULFID 464 476
FT DISULFID 506 543
FT DISULFID 510 548
FT DISULFID 521 533
FT DISULFID 553 564
FT DISULFID 558 574
FT DISULFID 577 588
FT DISULFID 594 610
FT DISULFID 601 619
FT DISULFID 622 646
FT DISULFID 652 665
FT DISULFID 659 678
FT DISULFID 680 691
FT DISULFID 707 715
HSP: P00740; 1EDM.
Genew; HGNC:11786; THBS2.
GO: 0008201; F: heparin binding; TAS.
MIM: 188061; -.
InterPro; IPR001881; EGF Ca.
InterPro; IPR006209; EGF like.
InterPro; IPR006210; IEGF.
InterPro; IPR000884; TSP1.
InterPro; IPR008085; TSP 1.
InterPro; IPR003367; tsp 3.
InterPro; IPR008089; TSPC.
InterPro; IPR003129; TSPN.
InterPro; IPR001007; WFC_C.
Pfam; PF00008; EGF; 2.
Pfam; PF00090; tsp_1; 3.
Pfam; PF02412; tsp_3; 13.
Pfam; PF05735; TSPN; 1.
Pfam; PF02210; TSPN; 1.
Pfam; PF00093; wvc; 1.
PRINTS; PRO1705; TSP1REPEAT.
SMART; SM00181; EGF; 3.
SMART; SM00209; TSP1; 3.
SMART; SM00210; TSPN; 1.
SMART; SM00214; WVC; 1.
PROSITE; PS00022; EGF 1; FALSE_NEG.
PROSITE; PS01186; EGF_2; 1.
PROSITE; PS00026; EGF_3; 2.
PROSITE; PS00092; TSP1; 3.
PROSITE; PS01208; WFC 1; 1.
PROSITE; PS01084; WFC_2; 1.
KW Glycoprotein; Cell adhesion; Calcium-binding; Heparin-binding; Repeat;
KW EGF-like domain; Signal.
FT SIGNAL 1 18
FT CHAIN 19 1172
FT DOMAIN 19 215
FT DOMAIN 19 232
FT DOMAIN 318 375
FT DOMAIN 381 431
FT DOMAIN 437 492
FT DOMAIN 494 549
FT DOMAIN 549 589
FT DOMAIN 590 647
FT DOMAIN 648 692
FT DOMAIN 725 760
FT DOMAIN 761 783
FT DOMAIN 784 819
FT DOMAIN 820 842
FT DOMAIN 843 880
FT DOMAIN 881 916
FT DOMAIN 917 952
FT DOMAIN 953 1172
FT SITE 928 930
FT DISULFID 266 266
FT DISULFID 270 270
FT DISULFID 393 425
FT DISULFID 397 430
FT DISULFID 408 415
FT DISULFID 449 486
FT DISULFID 453 491
FT DISULFID 464 476
FT DISULFID 506 543
FT DISULFID 510 548
FT DISULFID 521 533
FT DISULFID 553 564
FT DISULFID 558 574
FT DISULFID 577 588
FT DISULFID 594 610
FT DISULFID 601 619
FT DISULFID 622 646
FT DISULFID 652 665
FT DISULFID 659 678
FT DISULFID 680 691
FT DISULFID 707 715

```


laminin and type V collagen.

-!- SUBUNIT: Homotrimer; disulfide-linked.

-!- SIMILARITY: Belongs to the thrombospondin family.

-!- SIMILARITY: Contains 1 WFCC domain.

-!- SIMILARITY: Contains 3 EGF-like domains.

-!- SIMILARITY: Contains 3 TSP type-1 domains.

-!- SIMILARITY: Contains 7 TSP type-3 domains.

-!- SIMILARITY: Contains 1 TSP N-terminal (TSPN) domain.

This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See <http://www.isb-sib.ch/announce/> or send an email to license@isb-sib.ch).

EMBL; L07803; AAA33064.1; -.

EMBL; M64866; AAA40432.1; -.

PIR; A42587; A42587.

HSSP; P00740; 1EDM.

MGI; 98738; Thbs2.

InterPro; IPR001881; EGF_Ca.

InterPro; IPR006209; EGF like.

InterPro; IPR006210; IEGF.

InterPro; IPR008884; TSP1.

InterPro; IPR008885; TSP 1.

InterPro; IPR003387; TSP 3.

InterPro; IPR008859; TSPC.

InterPro; IPR003129; TSPC.

InterPro; IPR001007; VWF_C.

Pfam; PF00008; EGF; 2.

Pfam; PF00090; TSP 1; 3.

Pfam; PF02412; TSP 3; 13.

Pfam; PF05735; TSPC; 1.

Pfam; PF02210; TSPN; 1.

Pfam; PF00093; vwc; 1.

PRINTS; PR01705; TSP1REPEAT.

SMART; SM00181; EGF; 3.

SMART; SM00209; TSP1; 3.

SMART; SM00210; TSPN; 1.

SMART; SM00214; VWC; 1.

PROSITE; PS00022; EGF 1; FALSE_NEG.

PROSITE; PS01186; EGF 2; 1.

PROSITE; PS00026; EGF 3; 2.

PROSITE; PS00092; TSP1; 3.

PROSITE; PS01208; VWF_C; 1.

PROSITE; PS0184; VWF_C; 2; 1.

Glycoprotein; Cell adhesion; Calcium-binding; Heparin-binding; Repeat; EGF-like domain; Signal.

POTENTIAL.

SIGNAL 1 18

CHAIN 19 1172

THROMBOSPONDIN 2.

TSP N-TERMINAL.

HEPARIN-BINDING (POTENTIAL).

WFCC.

TSP TYPE-1 1.

TSP TYPE-1 2.

TSP TYPE-1 3.

EGF-LIKE 1.

EGF-LIKE 2.

EGF-LIKE 3.

TSP TYPE-3 1.

TSP TYPE-3 2.

TSP TYPE-3 3.

TSP TYPE-3 4.

TSP TYPE-3 5.

TSP TYPE-3 6.

TSP TYPE-3 7.

C-TERMINAL.

CELL ATTACHMENT SITE (POTENTIAL).

INTERCHAIN (PROBABLE).

INTERCHAIN (PROBABLE).

BY SIMILARITY.

FT DISULFID 397 430 BY SIMILARITY.

FT DISULFID 408 415 BY SIMILARITY.

FT DISULFID 449 486 BY SIMILARITY.

FT DISULFID 453 491 BY SIMILARITY.

FT DISULFID 464 476 BY SIMILARITY.

FT DISULFID 506 543 BY SIMILARITY.

FT DISULFID 510 548 BY SIMILARITY.

FT DISULFID 521 533 BY SIMILARITY.

FT DISULFID 553 564 BY SIMILARITY.

FT DISULFID 558 574 BY SIMILARITY.

FT DISULFID 577 588 BY SIMILARITY.

FT DISULFID 594 610 BY SIMILARITY.

FT DISULFID 601 619 BY SIMILARITY.

FT DISULFID 622 646 BY SIMILARITY.

FT DISULFID 652 665 BY SIMILARITY.

FT DISULFID 659 678 BY SIMILARITY.

FT DISULFID 680 691 BY SIMILARITY.

FT DISULFID 707 715 BY SIMILARITY.

FT DISULFID 720 740 BY SIMILARITY.

FT DISULFID 756 776 BY SIMILARITY.

FT DISULFID 779 799 BY SIMILARITY.

FT DISULFID 815 835 BY SIMILARITY.

FT DISULFID 838 858 BY SIMILARITY.

FT DISULFID 876 896 BY SIMILARITY.

FT DISULFID 912 932 BY SIMILARITY.

FT DISULFID 948 1169 BY SIMILARITY.

FT CARBOHYD 151 151 N-LINKED (GLCNAC. .) (POTENTIAL).

FT CARBOHYD 316 316 N-LINKED (GLCNAC. .) (POTENTIAL).

FT CARBOHYD 330 330 N-LINKED (GLCNAC. .) (POTENTIAL).

FT CARBOHYD 457 457 N-LINKED (GLCNAC. .) (POTENTIAL).

FT CARBOHYD 584 584 N-LINKED (GLCNAC. .) (POTENTIAL).

FT CARBOHYD 710 720 N-LINKED (GLCNAC. .) (POTENTIAL).

FT CARBOHYD 1069 1069 N-LINKED (GLCNAC. .) (POTENTIAL).

SQ SEQUENCE 1172 AA; 129911 MW; 7CE8E4E859822AB CRC64;

Query Match 5.7%; Score 273.5; DB 1; Length 1172;

Best Local Similarity 36.1%; Pred. No. 7.9e-13;

Matches 57; Conservative 22; Mismatches 70; Indels 9; Gaps 4;

QY 209 ROARLADTANYTCVAKNIVARRRS-ASAAVIVYVNGSGWSTWTSVCSACGRGWQKRSR 267

Db 403 QRGSCDVTSTNCLGPSIQTRTCSLQKCDTRIRONGWHSWSSCSVTGCVGNVTRIR 462

QY 268 SCTNPAPLNGGAFCEGQNVHDTVSLVSDGWSWPSKWSACGLDCT---HWRSRECS 324

Db 463 LCNSPVPQMGKCKNGSGRETKPCQDPCPIDGRNPSWSPNSACTVTCAGGIRERSVCN 522

QY 325 DPAPRNGRECOG--TD---LDTRNCTSLCVHVSASGP 357

Db 523 SPEPYGGKDCVGDVTEHQMCKNRSCPIDGCLSNPCFP 560

RESULT 4

SM5A MOUSE

ID SM5A MOUSE STANDARD; PRT; 1077 AA.

AC Q6217.

DT 30-MAY-2000 (Rel. 39, Created)

DT 30-MAY-2000 (Rel. 39, Last sequence update)

DT 28-FEB-2003 (Rel. 41, Last annotation update)

DE Semaphorin 5A precursor (Semaphorin F) (Sema F).

GN SEMA5A OR SEMAP OR SEMF.

OS Mus musculus (Mouse).

OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.

OX NCBI_TaxID=10090;

RN [1]

RP SEQUENCE FROM N.A.

RC STRAIN=NMRI;

EX MEDLINE=96414430; PubMed=8817451;

RA Adams R.H., Betz H., Püeschel A.W.;

RT "A novel class of murine semaphorins with homology to thrombospondin is differentially expressed during early embryogenesis.";

RL Mech. Dev. 57:33-45(1996).


```

FT DOMAIN 784 839 TSP TYPE-1 5.
FT DOMAIN 841 896 TSP TYPE-1 6.
FT DOMAIN 897 944 TSP TYPE-1 7.
FT CARBOHYD 142 142 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 168 168 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 227 227 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 277 277 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 323 323 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 367 367 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 437 437 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 536 536 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 591 591 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 717 717 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 933 933 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CONFLICT 56 56 A -> V (IN REF. 2).
FT CONFLICT 149 149 A -> T (IN REF. 2).
FT CONFLICT 382 382 V -> M (IN REF. 2).
FT CONFLICT 494 494 S -> R (IN REF. 2).
SQ SEQUENCE 1074 AA; 120570 MW; EE3DB763CBE29407 CRC64;

Query Match
Best Local Similarity 5.7%; Score 270.5; DB 1; Length 1074;
Matches 50; Conservative 13; Mismatches 52; Indels 3; Gaps 1;

QY 241 VNGGWSWTWESVCSACGRGMRKRSCTNPAPLNGGAFCEQNVHDTVSSLLVSDG 300
Db 783 VNGANSANTWGCSDRCGRGIRNRKVCNPPFKYGGMPCGLGPSLEYQECNTLPCPDG 842
QY 301 SWSPPSKWACGLDC---THWSRECSDFAPRNGBECQGTDLDRNCTSDLCVHSAS 355
Db 843 VNSCSPWTKCATCGGHHYMRTRSCNPAPAYGGDCLGLHTEALCNTQPCPWS 900

RESULT 6
TSP2_BOVIN STANDARD; PRT; 1170 AA.
AC Q95416; Q28180;
DT 01-NOV-1997 (Rel. 35, Created)
DT 15-JUL-1999 (Rel. 38, Last sequence update)
DT 28-FEB-2003 (Rel. 41, Last annotation update)
DE Thrombospondin 2 precursor (Corticotropin-induced secreted protein) (CISP).
GN THB2 OR TSP2 OR TSP-2.
OS Bos taurus (Bovine).
OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Cetartiodactyla; Ruminantia; Pecora; Bovidae;
OC Bovidae; Bovinae; Bos.
OX NCBI_TaxID=9913;
RN [1]
RP SEQUENCE FROM N.A.
RA Danik M., Chinn A., Lafeuillade M., Keramidas M., Aquesse-Germon S.,
RA Penhoat A., Chen H., Mosher D., Chambaz E.M., Feige J.J.;
RA Submitted (MAR-1998) to the EMBL/GenBank/DBJ databases.
RL [2]
RP SEQUENCE OF 1-522 FROM N.A.
RX MEDLINE=96331130; PubMed=8698834;
RA Lafeuillade B., Pellerin S., Keramidas M., Danik M., Chambaz E.M.,
RA Feige J.J.;
RT "Opposite regulation of thrombospondin-1 and corticotropin-induced secreted protein/thrombospondin-2 expression by adrenocorticotrophic hormone in adrenocortical cells.";
RT hormone in adrenocortical cells.";
RL J. Cell. Physiol. 167:164-172(1996).
RN [3]
RP SEQUENCE OF 318-831 FROM N.A.
RC TISSUE=Arctic endothelium;
RA Zafar R.S., Moll Y.D., Womack J.F., Walz D.A.;
RT "Cloning and sequencing of bovine thrombospondin stimulatory effect of TGF-beta.";
RT Submitted (MAY-1995) to the EMBL/GenBank/DBJ databases.
CC -!- FUNCTION: Adhesive glycoprotein that mediates cell-to-cell and cell-to-matrix interactions. Can bind to fibrinogen, fibronectin, laminin and type V collagen.
CC -!- SUBUNIT: Homotrimer; disulfide-linked.

```

```

CC -!- SIMILARITY: Belongs to the thrombospondin family.
CC -!- SIMILARITY: Contains 1 WFCC domain.
CC -!- SIMILARITY: Contains 3 EGF-like domains.
CC -!- SIMILARITY: Contains 7 TSP type-1 domains.
CC -!- SIMILARITY: Contains 7 TSP type-3 domains.
CC -!- SIMILARITY: Contains 1 TSP N-terminal (TSPN) domain.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See http://www.isb-sib.ch/announce/ or send an email to license@isb-sib.ch).
CC -----
CC EMBL; X96540; CAA65385.1; -.
CC EMBL; X87620; CAA60952.1; -.
CC HSSP; P00740; IEDM.
DR InterPro; IPR001881; EGF_Ca.
DR InterPro; IPR006209; EGF_like.
DR InterPro; IPR006210; IEGF.
DR InterPro; IPR000884; TSP1.
DR InterPro; IPR008085; TSP1.
DR InterPro; IPR003367; tsp_3.
DR InterPro; IPR008859; TSPC.
DR InterPro; IPR003129; TSPN.
DR InterPro; IPR001007; WFCC.
DR Pfam; PF00008; EGF; 1.
DR Pfam; PF00090; tsp_1; 3.
DR Pfam; PF02412; tsp_3; 13.
DR Pfam; PF05735; TSPC; 1.
DR Pfam; PF02210; TSPN; 1.
DR Pfam; PF00093; WVC; 1.
DR PRINTS; PRO1705; TSP1REPEAT.
DR SMART; SM00181; EGF; 3.
DR SMART; SM00209; TSP1; 3.
DR SMART; SM00210; TSPN; 1.
DR SMART; SM00214; WVC; 1.
DR PROSITE; PS00022; EGF_1; FALSE_NEG.
DR PROSITE; PS01186; EGF_2; 1.
DR PROSITE; PS00026; EGF_3; 2.
DR PROSITE; PS00092; TSP1; 3.
DR PROSITE; PS01208; WFCC_1; 1.
DR PROSITE; PS01184; WFCC_2; 1.
KW Glycoprotein; Cell adhesion; Calcium-binding; Heparin-binding; Repeat;
KW EGF-like domain; Signal.
FT SIGNAL 1 18
FT CHAIN 19 1170
FT DOMAIN 19 215 THROMBOSPONDIN 2.
FT DOMAIN 19 232 TSP N-TERMINAL.
FT DOMAIN 19 232 HEPARIN-BINDING (POTENTIAL).
FT DOMAIN 318 375 WFCC.
FT DOMAIN 379 429 TSP TYPE-1 1.
FT DOMAIN 435 490 TSP TYPE-1 2.
FT DOMAIN 492 547 TSP TYPE-1 3.
FT DOMAIN 547 597 EGF-LIKE 1.
FT DOMAIN 588 645 EGF-LIKE 2.
FT DOMAIN 646 690 EGF-LIKE 3.
FT DOMAIN 723 758 TSP TYPE-3 1.
FT DOMAIN 759 781 TSP TYPE-3 2.
FT DOMAIN 782 817 TSP TYPE-3 3.
FT DOMAIN 818 840 TSP TYPE-3 4.
FT DOMAIN 841 878 TSP TYPE-3 5.
FT DOMAIN 879 914 TSP TYPE-3 6.
FT DOMAIN 915 950 TSP TYPE-3 7.
FT DOMAIN 951 1170 C-TERMINAL.
FT SITE 926 928 CELL ATTACHMENT SITE (POTENTIAL).
FT DISULFID 266 266 INTERCHAIN (PROBABLE).
FT DISULFID 270 270 INTERCHAIN (PROBABLE).
FT DISULFID 391 423 BY SIMILARITY.
FT DISULFID 395 428 BY SIMILARITY.
FT DISULFID 406 413 BY SIMILARITY.
FT DISULFID 447 484 BY SIMILARITY.
FT DISULFID 451 489 BY SIMILARITY.

```

```

FT DISULFID 462 474 BY SIMILARITY.
FT DISULFID 504 541 BY SIMILARITY.
FT DISULFID 508 546 BY SIMILARITY.
FT DISULFID 519 531 BY SIMILARITY.
FT DISULFID 551 562 BY SIMILARITY.
FT DISULFID 556 572 BY SIMILARITY.
FT DISULFID 575 586 BY SIMILARITY.
FT DISULFID 592 608 BY SIMILARITY.
FT DISULFID 599 617 BY SIMILARITY.
FT DISULFID 620 644 BY SIMILARITY.
FT DISULFID 650 663 BY SIMILARITY.
FT DISULFID 676 676 BY SIMILARITY.
FT DISULFID 678 689 BY SIMILARITY.
FT DISULFID 705 713 BY SIMILARITY.
FT DISULFID 718 738 BY SIMILARITY.
FT DISULFID 754 774 BY SIMILARITY.
FT DISULFID 777 797 BY SIMILARITY.
FT DISULFID 813 833 BY SIMILARITY.
FT DISULFID 836 856 BY SIMILARITY.
FT DISULFID 874 894 BY SIMILARITY.
FT DISULFID 910 930 BY SIMILARITY.
FT DISULFID 946 1167 BY SIMILARITY.
FT CARBOHYD 151 151 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 316 316 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 330 330 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 455 455 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 582 582 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 708 708 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 936 936 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 1067 1067 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CONFLICT 535 535 A -> V (IN REF. 3).
FT CONFLICT 748 748 S -> T (IN REF. 3).
SQ SEQUENCE 1170 AA; 129862 MW; 9CF1EF5B89A051 CRC64;

Query Match
Best Local Similarity 35.6%; Pred. No. 1.8e-12;
Matches 52; Conservative 21; Mismatches 69; Indels 4; Gaps 2;

QY 209 RQRLADTANYTCVAKNIIVARRSASAA-VIVVINGGWSWTWTSVCSASCGRWQKR 267
DB 401 QGRSCDVTSNTCLGSIQTRACSLGRCDHRIRQDGGWSHWSFWSVSCVTCGVGNVTR 460
QY 268 SCTNPAPLNGGAFCEQNVHRTVSSLVSDGWSFKWACGLDCT--HWRSECS 324
DB 461 LCNSPVPMGGRSCKSGRETKACQPCPDVGRWSPFWSACTVTCAGGIRERTRVCN 520
QY 325 DPAPRNGBECQGTDLDTNCTSDLC 350
DB 521 SPEPQHGKDCVCGAKEQCMCRKSC 546

RESULT 7
BAI3 HUMAN STANDARD; PRT; 1522 AA.
AC O60297;
DT 16-OCT-2001 (Rel. 40, Created)
DT 16-OCT-2001 (Rel. 40, Last sequence update)
DT 28-FEB-2003 (Rel. 41, Last annotation update)
DE Brain-specific angiogenesis inhibitor 3 precursor.
GN BAI3 OR KIAA0550.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=Fetal brain;
RX MEDLINE=98194217; PubMed=9533023;
RA Shiratsuchi T., Nishimori H., Ichise H., Nakamura Y., Tokino T.;
RT "Cloning and characterization of BAI2 and BAI3, novel genes homologous
to brain-specific angiogenesis inhibitor 1 (BAI1).";
RL Cytogenet. Cell Genet. 79:103-108(1997).
RN [2]

```

```

RP SEQUENCE FROM N.A.
RC TISSUE=Brain;
RX MEDLINE=98290545; PubMed=9628581;
RA Nagase T., Ishikawa K.-I., Miyajima N., Tanaka A., Kotani H.,
RA Nomura N., Ohara O.;
RT "Prediction of the coding sequences of unidentified human genes. IX.
The complete sequences of 100 new cDNA clones from brain which can
code for large proteins in vitro.";
RL DNA Res. 5:31-39(1998).
RN [3]
RP SEQUENCE FROM N.A., AND REVISIONS TO 643-665 AND C-TERMINUS.
RX MEDLINE=22158633; PubMed=12168954;
RA Nakajima D., Okazaki N., Yamakawa H., Kikuno R., Ohara O., Nagase T.;
RT "Construction of expression-ready cDNA clones for KIAA genes: manual
curation of 330 KIAA cDNA clones.";
RL DNA Res. 9:99-106(2002).
CC -!- FUNCTION: MIGHT BE INVOLVED IN ANGIOGENESIS INHIBITION AND
SUPPRESSION OF GLOBLASTOMA.
CC -!- SUBCELLULAR LOCATION: Integral membrane protein.
CC -!- TISSUE SPECIFICITY: STRONGLY EXPRESSED IN BRAIN. ALSO DETECTED IN
HEART. REDUCED EXPRESSION IS OBSERVED IN SOME GLOBLASTOMA CELL
LINES.
CC -!- SIMILARITY: Belongs to family 2 of G-protein coupled receptors.
CC -!- SIMILARITY: Contains 1 CUB domain.
CC -!- SIMILARITY: Contains 4 TSP type-1 domains.
CC -!- SIMILARITY: Contains 1 GPS domain.
CC -----
CC THIS SWISS-PROT entry is copyrighted. It is produced through a collaboration
between the Swiss Institute of Bioinformatics and the EMBL outstation -
the European Bioinformatics Institute. There are no restrictions on its
use by non-profit institutions as long as its content is in no way
modified and this statement is not removed. Usage by and for commercial
entities requires a license agreement (See http://www.isb-sib.ch/announce/
or send an email to license@isb-sib.ch).
CC -----
DR EMBL; AB005299; BAA25363.1; -
DR EMBL; AB011122; BAA25476.2; ALT_INIT.
DR PIR; T00028; T00028
DR Genew; HGNC:945; BAI3.
DR MIM; 602684; -
DR InterPro; IPR000859; CUB.
DR InterPro; IPR000832; GPCR secretin.
DR InterPro; IPR001879; hormn_receptor.
DR InterPro; IPR000203; PKD_cys_rich.
DR InterPro; IPR000884; TSP1.
DR Pfam; PF00002; 7tm_2; 1.
DR Pfam; PF01825; GPS; 1.
DR Pfam; PF02793; HRM; 1.
DR Pfam; PF00090; tsp_1; 4.
DR SMART; SM00303; GPS; 1.
DR SMART; SM00008; Hormr; 1.
DR SMART; SM00209; TSP1; 4.
DR PROSITE; PS01180; CUB; 1.
DR PROSITE; PS0221; GPS; 1.
DR PROSITE; PS00649; G_PROTEIN_RECEP_F2_1; FALSE_NEG.
DR PROSITE; PS00650; G_PROTEIN_RECEP_F2_2; FALSE_NEG.
DR PROSITE; PS00227; G_PROTEIN_RECEP_F2_3; 1.
DR PROSITE; PS00261; G_PROTEIN_RECEP_F2_4; 1.
DR PROSITE; PS00092; TSP1; 4.
KW G-protein coupled receptor; Transmembrane; Glycoprotein; Signal;
Repeat.
FT SIGNAL 1 24 POTENTIAL.
FT CHAIN 25 1522 BRAIN-SPECIFIC ANGIOGENESIS INHIBITOR 3.
FT DOMAIN 25 880 EXTRACELLULAR (POTENTIAL).
FT TRANSMEM 881 901 1 (POTENTIAL).
FT DOMAIN 902 910 CYTOPLASMIC (POTENTIAL).
FT TRANSMEM 911 931 2 (POTENTIAL).
FT DOMAIN 932 939 EXTRACELLULAR (POTENTIAL).
FT TRANSMEM 940 960 3 (POTENTIAL).
FT DOMAIN 961 981 CYTOPLASMIC (POTENTIAL).
FT TRANSMEM 982 1002 4 (POTENTIAL).
FT DOMAIN 1003 1023 EXTRACELLULAR (POTENTIAL).
FT TRANSMEM 1024 1044 5 (POTENTIAL).

```

```

FT DOMAIN 1045 1098 CYTOPLASMIC (POTENTIAL).
FT TRANSMEM 1099 1119 6 (POTENTIAL).
FT DOMAIN 1120 1125 EXTRACELLULAR (POTENTIAL).
FT TRANSMEM 1126 1146 7 (POTENTIAL).
FT DOMAIN 1147 1522 CYTOPLASMIC (POTENTIAL).
FT FT DOMAIN 30 159 CUB.
FT FT DOMAIN 291 343 TSP TYPE-1 1.
FT FT DOMAIN 345 398 TSP TYPE-1 2.
FT FT DOMAIN 400 453 TSP TYPE-1 3.
FT FT DOMAIN 455 508 TSP TYPE-1 4.
FT FT DOMAIN 816 868 GPS.
FT FT DOMAIN 942 945 POLY-THR.
FT FT DOMAIN 1173 1176 POLY-SER.
FT FT CARBOHYD 51 51 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT FT CARBOHYD 54 54 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT FT CARBOHYD 82 82 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT FT CARBOHYD 105 105 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT FT CARBOHYD 241 241 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT FT CARBOHYD 337 337 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT FT CARBOHYD 418 418 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT FT CARBOHYD 540 540 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT FT CARBOHYD 625 625 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT FT CARBOHYD 779 779 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT FT CARBOHYD 812 812 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT FT CARBOHYD 828 828 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT FT CARBOHYD 937 937 N-LINKED (GLCNAC. . .) (POTENTIAL).
SQ SEQUENCE 1522 AA; 171490 MW; D22D0A5D4BB62502 CRC64;

Query Match 5.6%; Score 267.5; DB 1; Length 1522;
Best Local Similarity 37.7%; Pred. No. 3.3e-12;
Matches 55; Conservative 22; Mismatches 54; Indels 15; Gaps 5;

QY 220 TCVA-----KNIVARRRSAAVIVVNGGWSTWTSVCSACGKQKRSCTNPA 273
Db 317 TCVPSTGTCGSLRSLRVNNTALCPVGVWESWPSWLSCTPCGQGTQRTSCT--P 374
QY 274 PLNGGAFCGQNVHRTVSLVSDGWSWPSWKSACGLDC---THWRSRECSDEAPRN 330
Db 375 PQVGGRCFCEGPETHHKPCNIALCPVGVWESWPSWLSCTPCGQGTQRTSCT--AAAH 432
QY 331 GSEECQGTDLTRNCTSDLCVHSASG 356
Db 433 GSEECRGPWAESRECYNPEC--TANG 456

RESULT 8
TSPL_XENLA
ID TSPL_XENLA STANDARD; PRT; 1173 AA.
AC P35448;
DT 01-JUN-1994 (Rel. 29, Created)
DT 01-JUN-1994 (Rel. 29, Last sequence update)
DE 28-FEB-2003 (Rel. 41, Last annotation update)
DE Thrombospondin 1 precursor.
GN THBS1 OR TSPL.
OS Xenopus laevis (African clawed frog).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Amphibia; Batrachia; Anura; Mesobatrachia; Pipidea; Pipidae;
OC Xenopodinae; Xenopus.
OX NCBI_TaxID=8355;
RN [1]_TaxID=8355;
RP SEQUENCE FROM N.A.
RA Urry L.A., Ramos J., Duquette M., Desimone D.W., Lawler J.;
RL Submitted (XXX-1993) to the EMBL/GenBank/DBJ databases.
CC -!- FUNCTION: Adhesive glycoprotein that mediates cell-to-cell and
CC cell-to-matrix interactions. Can bind to fibrinogen, fibronectin,
CC laminin, type V collagen and integrins alpha-V/beta-1, alpha-
CC V/beta-3 and alpha-11b/beta-3 (by similarity).
CC -!- SUBUNIT: Homotrimer; disulfide-linked.
CC -!- SIMILARITY: Belongs to the thrombospondin family.
CC -!- SIMILARITY: Contains 1 VWFC domain.
CC -!- SIMILARITY: Contains 3 EGF-like domains.
CC -!- SIMILARITY: Contains 7 TSP type-1 domains.
CC -!- SIMILARITY: Contains 7 TSP type-3 domains.

```

```

CC -!- SIMILARITY: Contains 1 TSP N-terminal (TSPN) domain.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
CC EMBL; L04278; ; NOT ANNOTATED_CDS.
CC HSSP; P00740; 1EDM.
CC InterPro; IPR001881; EGF_Ca.
CC InterPro; IPR006209; EGF_Like.
CC InterPro; IPR006210; IEGF.
CC InterPro; IPR000884; TSP1.
CC InterPro; IPR008085; TSP 1.
CC InterPro; IPR003367; tsp 3.
CC InterPro; IPR008859; TSPC.
CC InterPro; IPR003129; TSPN.
CC InterPro; IPR001007; VWFC_C.
CC Pfam; PF00008; EGF; 2.
CC Pfam; PF00090; tsp 1; 3.
CC Pfam; PF02412; tsp 3; 13.
CC Pfam; PF05735; TSPC; 1.
CC Pfam; PF02210; TSPN; 1.
CC Pfam; PF00093; VWFC; 1.
CC PRINTS; PR01705; TSP1REPEAT.
CC SMART; SM00181; EGF; 2.
CC SMART; SM00209; TSP1; 3.
CC SMART; SM00210; TSPN; 1.
CC SMART; SM00214; VWFC; 1.
CC PROSITE; PS00022; EGF_1; FALSE_NEG.
CC PROSITE; PS01186; EGF_2; 1.
CC PROSITE; PS00026; EGF_3; 2.
CC PROSITE; PS00092; TSP1; 3.
CC PROSITE; PS01208; VWFC 1; 1.
CC PROSITE; PS01184; VWFC 2; 1.
CC Glycoprotein; Cell adhesion; Calcium-binding; Heparin-binding; Repeat;
CC EGF-like domain; signal.
CC SIGNAL 1 22
FT CHAIN 23 1173 THROMBOSPONDIN 1.
FT DOMAIN 23 235 HEPARIN-BINDING (POTENTIAL).
FT DOMAIN 23 224 TSP N-TERMINAL.
FT DOMAIN 319 376 VWFC.
FT DOMAIN 382 432 TSP TYPE-1 1.
FT DOMAIN 438 493 TSP TYPE-1 2.
FT DOMAIN 495 550 TSP TYPE-1 3.
FT DOMAIN 550 590 EGF-LIKE 1.
FT DOMAIN 591 648 EGF-LIKE 2, CALCIUM-BINDING (POTENTIAL).
FT DOMAIN 649 693 EGF-LIKE 3.
FT DOMAIN 726 761 TSP TYPE-3 1.
FT DOMAIN 762 784 TSP TYPE-3 2.
FT DOMAIN 785 820 TSP TYPE-3 3.
FT DOMAIN 821 843 TSP TYPE-3 4.
FT DOMAIN 844 881 TSP TYPE-3 5.
FT DOMAIN 882 917 TSP TYPE-3 6.
FT DOMAIN 918 953 TSP TYPE-3 7.
FT DOMAIN 954 1173 C-TERMINAL.
FT SITE 929 931 CELL ATTACHMENT SITE (POTENTIAL).
FT DISULFID 394 426 BY SIMILARITY.
FT DISULFID 398 431 BY SIMILARITY.
FT DISULFID 409 416 BY SIMILARITY.
FT DISULFID 450 487 BY SIMILARITY.
FT DISULFID 454 492 BY SIMILARITY.
FT DISULFID 465 477 BY SIMILARITY.
FT DISULFID 507 544 BY SIMILARITY.
FT DISULFID 511 549 BY SIMILARITY.
FT DISULFID 522 534 BY SIMILARITY.
FT DISULFID 554 565 BY SIMILARITY.
FT DISULFID 559 575 BY SIMILARITY.
FT DISULFID 578 589 BY SIMILARITY.
FT DISULFID 595 611 BY SIMILARITY.

```

```

FT DISULFID 602 BY SIMILARITY.
FT DISULFID 623 BY SIMILARITY.
FT DISULFID 653 BY SIMILARITY.
FT DISULFID 660 BY SIMILARITY.
FT DISULFID 681 BY SIMILARITY.
FT DISULFID 708 BY SIMILARITY.
FT DISULFID 721 BY SIMILARITY.
FT DISULFID 757 BY SIMILARITY.
FT DISULFID 780 BY SIMILARITY.
FT DISULFID 800 BY SIMILARITY.
FT DISULFID 816 BY SIMILARITY.
FT DISULFID 839 BY SIMILARITY.
FT DISULFID 877 BY SIMILARITY.
FT DISULFID 913 BY SIMILARITY.
FT DISULFID 949 BY SIMILARITY.
FT CARBOHYD 155 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 158 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 250 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 363 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 705 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 711 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 1070 N-LINKED (GLCNAC. .) (POTENTIAL).
SQ SEQUENCE 1173 AA; 130019 MW; A3P036D6516C0F24 CRC64;

Query Match 5.6%; Score 266.5; DB 1; Length 1173;
Best Local Similarity 22.5%; Pred. No. 2.7e-12;
Matches 86; Conservative 54; Mismatches 147; Indels 95; Gaps 14;

QY 11 LIGIVLAAMLRSG-----AQQSTVANVPYGANPDLLPHFLPEPEVDIVVKNKPVLLVC- 65
Db 221 VFGTTLEAILRNKGLSMNSVITLDPNVPNGSPAIRTYIGH-----KTKDLOAVCG 273

QY 66 -KAVPATQIFKNGEVRQVDHVIERSDTGSGEPTMEVRINVSQVKEVFGLEEYWC 124
Db 274 FSCDDLSKLFAMKG-----LRLVTLKQVTKETEKNELIA 311

QY 125 QCVAMSSGTTKSKAYIRIARLNKFEQ-----EPLAKEVSLEQGIVLPC 170
Db 312 QIV-----TRTPGVCLHNGVLHKNRDEWTDSCTECTCONSATICRKVSCP---LMFC 361

QY 171 RP-----IPAEVWELRNEDLPSPDPNVYITREHSLVVRQARLADTANY 219
Db 362 TNATIPDGBCCPCWPSDSADDDSPSWDTPCS-----VTCGHG-IQQRGSCDSLNN 414

QY 220 TCVAKNIVAR-----RRSAAVIVYVNGWSTWTSVCSACSGRGWQKRSCTN 271
Db 415 PCEGSSVQRTSCQIQCDCKRFQ-----DGGWSHSPSSCSVTCGSGQITRIKLNS 467

QY 272 PAPLNGGAFCEGQNVHRTVSSLLVSDGSPWSKWSACGLDC---THWRSRECSDPAP 328
Db 468 PVPQLNGKQCEGEGRENKPCQKDPCCPQNGQWGFWSLWDTCTVTCGGGQKRERLCNPXP 527

QY 329 RNGEECCQTDLDTRNCTSDLC 350
Db 528 QYEGKDCICEPTDSQICNKCQDC 549

RESULT 9
SM5B_MOUSE
ID SM5B_MOUSE STANDARD; PRT; 1093 AA.
AC Q60519;
DT 30-MAY-2000 (Rel. 39, Created)
DT 30-MAY-2000 (Rel. 39, Last sequence update)
DT 28-FEB-2003 (Rel. 41, Last annotation update)
DE Semaphorin 5B precursor (Semaphorin G) (Sema G).
GN Sema5B OR SEMAG OR SEMG.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=NMRI;
RX MEDLINE=96414430; PubMed=8817451;

```

```

RA Adams R.H., Betz H., Puschel A.W.;
RT "A novel class of murine semaphorins with homology to thrombospondin
RT is differentially expressed during early embryogenesis.";
RL Mech. Dev. 57:33-45(1996).
CC -!- FUNCTION: May act as positive axonal guidance cues.
CC -!- SUBCELLULAR LOCATION: Type I membrane protein.
CC -!- TISSUE SPECIFICITY: In adult, only detected in brain.
CC -!- DEVELOPMENTAL STAGE: Differentially expressed in embryonic and
CC adult tissues. Its abundance decreases from E10 to birth.
CC -!- SIMILARITY: Belongs to the semaphorin family.
CC -!- SIMILARITY: Contains 1 Sema domain.
CC -!- SIMILARITY: Contains 7 TSP type-1 domains.
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL; X97818; CAA66398.1; -.
DR MGD; MGI:107555; Sema5b.
DR InterPro; IPR003659; Plexin-like.
DR InterPro; IPR002165; Plexin_repeat.
DR InterPro; IPR001627; Sema.
DR InterPro; IPR000884; TSP1.
DR InterPro; IPR008085; TSP_1.
DR Pfam; PF01437; PSI; 1.
DR Pfam; PF01403; Sema; 1.
DR Pfam; PF00090; tsp 1; 5.
DR PRINTS; PR01705; TSP1REPEAT.
DR SMART; SM00423; PSI; 1.
DR SMART; SM00630; Sema; 1.
DR SMART; SM00209; TSP1; 4.
DR PROSITE; PS50092; TSP1; 5.
KW Signal; Transmembrane; Repeat; Multigene family; Neurogenesis;
KW Developmental protein; Glycoprotein.
FT SIGNAL 1 19
FT CHAIN 20 1093 SEMAPHORIN 5B.
FT DOMAIN 20 978 EXTRACELLULAR (POTENTIAL).
FT TRANSMEM 979 999 POTENTIAL.
FT DOMAIN 1000 1093 CYTOPLASMIC (POTENTIAL).
FT DOMAIN 236 518 SEMA.
FT DOMAIN 551 605 TSP TYPE-1 1.
FT DOMAIN 606 662 TSP TYPE-1 2.
FT DOMAIN 664 713 TSP TYPE-1 3.
FT DOMAIN 721 776 TSP TYPE-1 4.
FT DOMAIN 795 850 TSP TYPE-1 5.
FT DOMAIN 852 907 TSP TYPE-1 6.
FT DOMAIN 908 952 TSP TYPE-1 7.
FT CARBOHYD 59 59 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 95 95 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 157 157 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 178 178 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 287 287 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 333 333 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 378 378 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 532 532 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 539 539 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 547 547 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 602 602 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 728 728 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 944 944 N-LINKED (GLCNAC. .) (POTENTIAL).
SQ SEQUENCE 1093 AA; 120326 MW; 29ESC9B1E8108717 CRC64;

Query Match 5.4%; Score 258.5; DB 1; Length 1093;
Best Local Similarity 22.2%; Pred. No. 1e-11;
Matches 104; Conservative 31; Mismatches 139; Indels 195; Gaps 16;

QY 163 EQGIVLPCRPPEGIPPAEVEWLRNEDLPSPDPNVYITREHSLVVRQARLADTANYTCV 222
Db 737 EQRPRTCRAP-----LPDP-----HGLQFGKR---TETRTCP 767

```



```
Db 967 EATGCAGFNLLHLVATGISCFLGSLTAVLYLSQCQCQSQSESTL----- 1013
QY 409 STKPSKADNPHLLTIOPDLSTTTTQSGSLCPQDGPSP-KFQLTNGHLHLP 458
Db 1014 -VHPATPNHLH-----YKGGGTPKNEKVTPEFKTLNKNLIP 1050

RESULT 12
TSPI_MOUSE
ID TSPI_MOUSE STANDARD; PRT; 1170 AA.
AC P35441;
DT 01-JUN-1994 (Rel. 29, Created)
DT 01-JUN-1994 (Rel. 29, Last sequence update)
DT 28-FEB-2003 (Rel. 41, Last annotation update)
DE Thrombospondin 1 precursor.
GN THBS1 OR TSPI.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=92128941; PubMed=1774063;
RA Lawler J., Duquette M., Ferro P., Copeland N.G., Gilbert D.J.,
RA Jenkins N.A.;
RT "Characterization of the murine thrombospondin gene.";
RL Genomics 11:587-600(1991).
RN [2]
RP SEQUENCE FROM N.A.
RX MEDLINE=92147683; PubMed=1371115;
RA Laherty C.D., O'Rourke K., Wolf F.W., Katz R., Seldin M.F.,
RA Dixit V.M.;
RT "Characterization of mouse thrombospondin 2 sequence and expression
during cell growth and development.";
RL J. Biol. Chem. 267:3274-3281(1992).
RN [3]
RP SEQUENCE OF 1-490 FROM N.A.
RX MEDLINE=90375546; PubMed=2398070;
RA Bornstein P., Alfi D., Devarayalu S., Franson P., Li P.;
RT "Characterization of the mouse thrombospondin gene and evaluation of
the role of the first intron in human gene expression.";
RL J. Biol. Chem. 265:16691-16698(1990).
CC -!- FUNCTION: Adhesive glycoprotein that mediates cell-to-cell and
cell-to-matrix interactions. Can bind to fibrinogen, fibronectin,
laminin, type V collagen and integrins alpha-V/beta-1, alpha-
V/beta-3 and alpha-1ib/beta-3.
CC -!- SUBUNIT: Homotrimer; disulfide-linked.
CC -!- SIMILARITY: Belongs to the thrombospondin family.
CC -!- SIMILARITY: Contains 1 WFCC domain.
CC -!- SIMILARITY: Contains 3 EGF-like domains.
CC -!- SIMILARITY: Contains 3 TSP type-1 domains.
CC -!- SIMILARITY: Contains 7 TSP type-3 domains.
CC -!- SIMILARITY: Contains 1 TSP N-terminal (TSPN) domain.
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
between the Swiss Institute of Bioinformatics and the EMBL outstation -
the European Bioinformatics Institute. There are no restrictions on its
use by non-profit institutions as long as its content is in no way
modified and this statement is not removed. Usage by and for commercial
entities requires a license agreement (See http://www.isb-sib.ch/announce/
or send an email to license@isb-sib.ch).
CC
DR EMBL; M62470; AAA50611.1; -.
DR EMBL; M62450; AAA50611.1; JOINED.
DR EMBL; M62451; AAA50611.1; JOINED.
DR EMBL; M62452; AAA50611.1; JOINED.
DR EMBL; M62453; AAA50611.1; JOINED.
DR EMBL; M62454; AAA50611.1; JOINED.
DR EMBL; M62455; AAA50611.1; JOINED.
DR EMBL; M62456; AAA50611.1; JOINED.
DR EMBL; M62457; AAA50611.1; JOINED.
DR EMBL; M62458; AAA50611.1; JOINED.

DR EMBL; M62459; AAA50611.1; JOINED.
DR EMBL; M62460; AAA50611.1; JOINED.
DR EMBL; M62461; AAA50611.1; JOINED.
DR EMBL; M62462; AAA50611.1; JOINED.
DR EMBL; M62463; AAA50611.1; JOINED.
DR EMBL; M62464; AAA50611.1; JOINED.
DR EMBL; M62465; AAA50611.1; JOINED.
DR EMBL; M62466; AAA50611.1; JOINED.
DR EMBL; M62467; AAA50611.1; JOINED.
DR EMBL; M62468; AAA50611.1; JOINED.
DR EMBL; M62469; AAA50611.1; JOINED.
DR EMBL; M87276; AAA53063.1; -.
DR EMBL; J05605; AAA40431.1; -.
DR EMBL; J05605; AAA40431.1; JOINED.
DR PIR; A40558; AAA40431.1; JOINED.
DR MGD; MGI:98737; Thbs1.
DR GO; GO:0005615; C:extracellular space; IDA.
DR GO; GO:0016525; P:negative regulation of angiogenesis; IDA.
DR InterPro; IPR001881; EGF_Ca.
DR InterPro; IPR006209; EGF_like.
DR InterPro; IPR006210; IEGF.
DR InterPro; IPR000884; TSPI.
DR InterPro; IPR008085; TSP_1.
DR InterPro; IPR003367; tsp_3.
DR InterPro; IPR008859; TSPC.
DR InterPro; IPR003129; TSPN.
DR InterPro; IPR001007; VWF_C.
DR Pfam; PF00008; EGF; 2.
DR Pfam; PF00090; tsp_1; 3.
DR Pfam; PF02412; tsp_3; 13.
DR Pfam; PF05735; TSPC; 1.
DR Pfam; PF02210; TSPN; 1.
DR Pfam; PF00093; vwc; 1.
DR PRINTS; PR01705; TSPIREPEAT.
DR SMART; SM00181; EGF; 3.
DR SMART; SM00209; TSP1; 3.
DR SMART; SM00210; TSPN; 1.
DR SMART; SM00214; VNC; 1.
DR PROSITE; PS00022; EGF_1; FALSE_NEG.
DR PROSITE; PS01186; EGF_2; 1.
DR PROSITE; PS00026; EGF_3; 2.
DR PROSITE; PS00092; TSPI; 3.
DR PROSITE; PS01208; VWF_1; 1.
DR PROSITE; PS0184; VWF_2; 1.
KW Glycoprotein; Cell adhesion; Calcium-binding; Heparin-binding; Repeat;
KW EGF-like domain; Signal.
FT SIGNAL 1 18
FT CHAIN 19 1170
FT DOMAIN 19 232
FT DOMAIN 24 221
FT DOMAIN 316 373
FT DOMAIN 379 429
FT DOMAIN 435 490
FT DOMAIN 492 547
FT DOMAIN 549 587
FT DOMAIN 588 645
FT DOMAIN 646 690
FT DOMAIN 723 758
FT DOMAIN 759 781
FT DOMAIN 782 817
FT DOMAIN 818 840
FT DOMAIN 841 878
FT DOMAIN 879 914
FT DOMAIN 915 950
FT DOMAIN 951 1170
FT SITE 926 928
FT DISULFID 270 274
FT DISULFID 274 274
FT DISULFID 391 423
FT DISULFID 395 428
FT DISULFID 406 413
FT DISULFID 447 484
FT DISULFID 451 489

POTENTIAL.
THROMBOSPONDIN 1.
HEPARIN-BINDING (POTENTIAL).
TSP N-TERMINAL.
VWFC.
TSP TYPE-1 1.
TSP TYPE-1 2.
TSP TYPE-1 3.
EGF-LIKE 1.
EGF-LIKE 2.
EGF-LIKE 3.
TSP TYPE-3 1.
TSP TYPE-3 2.
TSP TYPE-3 3.
TSP TYPE-3 4.
TSP TYPE-3 5.
TSP TYPE-3 6.
TSP TYPE-3 7.
C-TERMINAL.
CELL ATTACHMENT SITE (POTENTIAL).
INTERCHAIN (PROBABLE).
INTERCHAIN (PROBABLE).
BY SIMILARITY.
BY SIMILARITY.
BY SIMILARITY.
BY SIMILARITY.
```


[illegible]

GenCore version 5.1.6
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: October 4, 2004, 18:21:41 ; Search time 140 Seconds
(without alignments)
2026.078 Million cell updates/sec

Title: US-09-970-944-2

Perfect score: 4787

Sequence: 1 MAVRPLWPLLVLAIAWL.....AVAGLQPDAGLFTVSEAEFC 899

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 1017041 seqs, 315518202 residues

Total number of hits satisfying chosen parameters: 1017041

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

SPTREMBL.25:*

- 1: sp_archaea:*
- 2: sp_bacteria:*
- 3: sp_fungi:*
- 4: sp_human:*
- 5: sp_invertebrate:*
- 6: sp_mammal:*
- 7: sp_mhc:*
- 8: sp_organelle:*
- 9: sp_phase:*
- 10: sp_plant:*
- 11: sp_rodent:*
- 12: sp_virus:*
- 13: sp_vertebrate:*
- 14: sp_unclassified:*
- 15: sp_rvirus:*
- 16: sp_bacteriap:*
- 17: sp_archaeap:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query %	Match Length	DB ID	Description
1	4592.5	95.9	898	11 Q8K1S4	Q8K1S4 mus musculus
2	4545.5	95.0	898	11 Q8G721	Q8G721 rattus norv
3	2819	58.9	544	4 Q96GP4	Q96GP4 homo sapien
4	2705.5	56.5	931	11 Q08747	Q08747 mus musculus
5	2686	56.1	950	11 Q8CD16	Q8CD16 mus musculus
6	2679.5	56.0	931	13 Q7T2Z5	Q7T2Z5 gallus gall
7	2673.5	55.8	931	4 Q95185	Q95185 mus musculus
8	2570	53.7	943	13 Q8UGT4	Q8UGT4 xenopus lae
9	2505.5	52.3	1008	11 Q8QY85	Q8QY85 mus musculus
10	2499	52.2	945	11 Q8K1S3	Q8K1S3 mus musculus
11	2497	52.2	945	11 Q08722	Q08722 rattus norv
12	2493	52.1	945	11 Q9D398	Q9D398 mus musculus
13	2486.5	51.9	934	4 Q8IZJ1	Q8IZJ1 homo sapien
14	2479	51.8	945	4 Q86SN3	Q86SN3 mus musculus
15	2144.5	44.8	956	11 Q8K1S2	Q8K1S2 mus musculus
16	2136	44.6	948	4 Q8WYP7	Q8WYP7 homo sapien

17	1597.5	33.4	597	4 Q8IUT0	Q8IUT0 homo sapien
18	1442	30.1	328	11 Q8OT71	Q8OT71 mus musculus
19	1223.5	25.6	554	4 Q8NIY2	Q8NIY2 homo sapien
20	971	20.3	366	4 Q9H9F3	Q9H9F3 homo sapien
21	961.5	20.1	1072	5 Q9NBL0	Q9NBL0 drosophila
22	956.5	20.0	1072	5 Q9V7B5	Q9V7B5 drosophila
23	953.5	19.9	947	5 Q26262	Q26262 caenorhabdi
24	950.5	19.9	947	5 Q44171	Q44171 caenorhabdi
25	694	14.5	199	13 Q9PVD5	Q9PVD5 petromyzon
26	552.5	11.5	351	4 Q8TF26	Q8TF26 homo sapien
27	366	7.6	2673	4 Q96SC3	Q96SC3 homo sapien
28	366	7.6	5636	4 Q96RW7	Q96RW7 homo sapien
29	311.5	6.5	325	5 Q811K1	Q811K1 drosophila
30	306	6.4	518	4 Q8IV45	Q8IV45 homo sapien
31	273.5	5.7	1172	11 Q8CG21	Q8CG21 mus musculus
32	273.5	5.7	1172	11 Q7TMT3	Q7TMT3 mus musculus
33	271	5.7	1582	11 Q8CGM0	Q8CGM0 mus musculus
34	270.5	5.7	1081	5 Q9U631	Q9U631 drosophila
35	269.5	5.6	1083	5 Q9VT70	Q9VT70 drosophila
36	269.5	5.6	1091	5 Q7YU67	Q7YU67 drosophila
37	267	5.6	1522	11 Q8OZFB	Q8OZFB mus musculus
38	263.5	5.5	1461	5 Q8MYA8	Q8MYA8 caenorhabdi
39	258.5	5.4	1122	11 Q7TT33	Q7TT33 mus musculus
40	258	5.4	1573	4 Q8NGW8	Q8NGW8 homo sapien
41	256	5.3	478	11 Q8BVE5	Q8BVE5 mus musculus
42	254.5	5.3	1560	11 Q8CGM1	Q8CGM1 mus musculus
43	249	5.2	885	6 Q9TTS5	Q9TTS5 bos taurus
44	249	5.2	5146	6 Q8SPM4	Q8SPM4 bos taurus
45	248	5.2	1171	11 Q8CGB2	Q8CGB2 mus musculus

ALIGNMENTS

RESULT 1

ID Q8K1S4 PRELIMINARY; PRT; 898 AA.
AC Q8K1S4;
DT 01-OCT-2002 (Tremblrel. 22, Created)
DT 01-OCT-2002 (Tremblrel. 22, Last sequence update)
DE 01-OCT-2003 (Tremblrel. 25, Last annotation update)
DE Netrin receptor Unc5h1.
GN UNC5H1.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
RN [1]
RP SEQUENCE FROM N.A.
RA Engelkamp D.;
RT "Cloning of three mouse unc-5 genes and their expression patterns at mid-gestation."
RL Submitted (MAY-2002) to the EMBL/GenBank/DBJ databases.
DR EMBL; AJ487852; CAD32250.1; -.
DR MGD; MGI:894682; Unc5h1.
DR GO; GO:0004872; F:receptor activity; IEA.
DR GO; GO:0007165; P:signal transduction; IEA.
DR InterPro; IPR000488; Death.
DR InterPro; IPR003599; IG.
DR InterPro; IPR007110; IG-like.
DR InterPro; IPR000884; TSP1.
DR InterPro; IPR008085; TSP1.
DR InterPro; IPR000906; ZUS.
DR Pfam; PF00531; Death; 1.
DR Pfam; PF00047; ig; 1.
DR Pfam; PF00090; tsp; 1.
DR Pfam; PF00791; ZUS; 1.
DR PRINTS; PR01705; TSP1REPEAT.
DR SMART; SM00005; DEATH; 1.
DR SMART; SM00409; IG; 1.
DR SMART; SM00209; TSP1; 2.
DR SMART; SM00218; ZUS; 1.
DR PROSITE; PS50835; IG_LIKE; 1.

DR PROSITE: PS50092; TSP1; 2.
KW Receptor. 898 AA; 98856 MW; 59F04BA2E196C1DB CRC64;
SQ SEQUENCE 898 AA; 98856 MW; 59F04BA2E196C1DB CRC64;
Query Match 95.9%; Score 4592.5; DB 11; Length 898;
Best Local Similarity 95.3%; Pred. No. 0;
Matches 858; Conservative 21; Mismatches 18; Indels 3; Gaps 3;
QY 1 MAVRPGMLPALLGIVLAAWLRSGGAQOSATVANVPVPGANPDLLPHELVPEDEVYIVKNKP 60
Db 1 MAVRPGMLPALLGIVLTAWLRSGGAQOSATVANVPVPGANPDLLPHELVPEDEVYIVKNKP 60
QY 61 VLLVCKAVPATQIFFKNGEWVQVDHVIERSDGSSEPTMEVRINVSQQKEKVFGL 120
Db 61 VLLVCKAVPATQIFFKNGEWVQVDHVIERSDGSSEPTMEVRINVSQQKEKVFGL 120
QY 121 EYWCQCVAMSSGTTKSKAYIRIARLRKNFEOEPLAKEVSLQGVILPCRPPEGIPPAE 180
Db 121 EYWCQCVAMSSGTTKSKAYIRIARLRKNFEOEPLAKEVSLQGVILPCRPPEGIPPAE 180
QY 181 VEWLNRNEDLVDSLDNNVYITREHSLVVRQARLADTANYTCVAKNIVARRRSAAVIVY 240
Db 181 VEWLNRNEDLVDSLDNNVYITREHSLVVRQARLADTANYTCVAKNIVARRRSAAVIVY 240
QY 241 VNGGWTWTEWSVCSACGRGWQKRSCTNPAPLNGGAFCEGQNV-QKTACATLCPVDG 300
Db 241 VNGGWTWTEWSVCSACGRGWQKRSCTNPAPLNGGAFCEGQNV-QKTACATLCPVDG 300
QY 301 SMSWSKWSACGLDCTHWRSECDPAPRNGEEOCGTDLTRNCTSDLCVHSASPEDV 360
Db 301 SMSWSKWSACGLDCTHWRSECDPAPRNGEEOCGTDLTRNCTSDLCVHSASPEDV 360
QY 361 ALYVGLIAVAVCLVLLVLLVLYCRKKEGLSDVADSSILTSFGFQVSIKPSKADNPHL 420
Db 360 ALYVGLIAVAVCLVLLVLLVLYCRKKEGLSDVADSSILTSFGFQVSIKPSKADNPHL 420
QY 421 LTIQPDLS-TTTTYQGSCLPRQDGPSPKFQLTNGHLLSPGGRHRLHSSPTSEAEFV 479
Db 421 LTIQPDLS-TTTTYQGSCLPRQDGPSPKFQLTNGHLLSPGGRHRLHSSPTSEAEFV 479
QY 480 SRLSTQNYFRSLPRGTSNMYTGFNFGRLMIPNTGISLLIPPDAPRGIYIYLTLH 539
Db 480 SRLSTQNYFRSLPRGTSNMYTGFNFGRLMIPNTGISLLIPPDAPRGIYIYLTLH 539
QY 540 KPEDVPLAGCOTLLSPIVSCGPPGVLLTRPVILAMDHCGEPPSDWSLRKKQCEGS 599
Db 540 KPEDVPLAGCOTLLSPIVSCGPPGVLLTRPVILAMDHCGEPPSDWSLRKKQCEGS 599
QY 600 WEQDVHLHGEAPSHLYYCOLEASACVFTQOLGRFALVGEALSVAALKLLLPAPVA 659
Db 600 WE-DVHLHGEAPSHLYYCOLEAGACVFTQOLGRFALVGEALSVAALKLLLPAPVA 658
QY 660 CTSLEYNIRVYCLHDTHDALKEVVQLEKQGGQIQEPRVLHPKDSYHNLRLSIHDVPS 719
Db 659 CTSLEYNIRVYCLHDTHDALKEVVQLEKQGGQIQEPRVLHPKDSYHNLRLSIHDVPS 718
QY 720 LMKSKLLVSYQETPFVHIWNGTQRYLHCTFTLBRVSPSTDLACKLWVQEGDGSFSI 779
Db 719 LMKSKLLVSYQETPFVHIWNGTQRYLHCTFTLBRVSPSTDLACKLWVQEGDGSFSI 778
QY 780 NFENITKDTFAELLALALEAGVPALVGPFAKIPFLIRQKIISLDPCCRRGADWTLAQ 839
Db 779 NFENITKDTFAELLALALEAGVPALVGPFAKIPFLIRQKIISLDPCCRRGADWTLAQ 838
QY 840 KLHLDHLSFFASKPSPTAMILNWRARHPNGLSOLAAVAGLQPDAGLPTVSEAE 899
Db 839 KLHLDHLSFFASKPSPTAMILNWRARHPNGLSOLAAVAGLQPDAGLPTVSEAE 898

RESULT 2
C08721
ID C08721
AC C08721;

DT 01-JUL-1997 (TrEMBLrel. 04, Created)
DT 01-JUL-1997 (TrEMBLrel. 04, Last sequence update)
DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
DE Transmembrane receptor UNC5H1.
OC Rattus norvegicus (Rat).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.
OX NCBI_TaxID=10116;
RN [1]
SEQUENCE FROM N.A.
RP TISSUE=Brain, and Ventral spinal cord;
RX MEDLINE=97271897; PubMed=9126742;
RA Leonardo E.D., Hinc L., Masu M., Keino-Masu K., Ackerman S.L.,
RA Tessier-Lavigne M.;
RT "vertebrate homologues of C. elegans UNC-5 are candidate netrin receptors.";
RL Nature 386:833-838(1997).
DR BMBL; U87305; AAB57678.1; -
DR GO; GO:0004872; F:receptor activity; IEA.
DR GO; GO:0007165; P:signal transduction; IEA.
DR InterPro; IPR000488; Death.
DR InterPro; IPR003599; IG.
DR InterPro; IPR007110; IG-like.
DR InterPro; IPR000884; TSP1.
DR InterPro; IPR008085; TSP1.
DR InterPro; IPR000906; ZUS.
DR Pfam; PF00531; death; 1.
DR Pfam; PF00047; ig; 1.
DR Pfam; PF00090; tsp1; 2.
DR Pfam; PF00791; ZUS; 1.
DR PRINTS; PR01705; TSP1REPEAT.
DR SMART; SM00005; DEATH; 1.
DR SMART; SM00409; IG; 1.
DR SMART; SM00209; TSP1; 2.
DR SMART; SM00218; ZUS; 1.
DR PROSITE; PS50835; IG LIKE; 1.
DR PROSITE; PS50092; TSP1; 2.
KW Receptor.
SQ SEQUENCE 898 AA; 98840 MW; 7A3CB9E7ACAL35 CRC64;
Query Match 95.0%; Score 4545.5; DB 11; Length 898;
Best Local Similarity 94.7%; Pred. No. 0;
Matches 852; Conservative 19; Mismatches 26; Indels 3; Gaps 3;
QY 1 MAVRPGMLPALLGIVLAAWLRSGGAQOSATVANVPVPGANPDLLPHELVPEDEVYIVKNKP 60
Db 1 MAVRPGMLPALLGIVLAAWLRSGGAQOSATVANVPVPGANPDLLPHELVPEDEVYIVKNKP 60
QY 61 VLLVCKAVPATQIFFKNGEWVQVDHVIERSDGSSEPTMEVRINVSQQKEKVFGL 120
Db 61 VLLVCKAVPATQIFFKNGEWVQVDHVIERSDGSSEPTMEVRINVSQQKEKVFGL 120
QY 121 EYWCQCVAMSSGTTKSKAYIRIARLRKNFEOEPLAKEVSLQGVILPCRPPEGIPPAE 180
Db 121 EYWCQCVAMSSGTTKSKAYIRIARLRKNFEOEPLAKEVSLQGVILPCRPPEGIPPAE 180
QY 181 VEWLNRNEDLVDSLDNNVYITREHSLVVRQARLADTANYTCVAKNIVARRRSAAVIVY 240
Db 181 VEWLNRNEDLVDSLDNNVYITREHSLVVRQARLADTANYTCVAKNIVARRRSAAVIVY 240
QY 241 VNGGWTWTEWSVCSACGRGWQKRSCTNPAPLNGGAFCEGQNV-QKTACATLCPVDG 300
Db 241 VNGGWTWTEWSVCSACGRGWQKRSCTNPAPLNGGAFCEGQNV-QKTACATLCPVDG 300
QY 301 SMSWSKWSACGLDCTHWRSECDPAPRNGEEOCGTDLTRNCTSDLCVHSASPEDV 360
Db 301 SMSWSKWSACGLDCTHWRSECDPAPRNGEEOCGTDLTRNCTSDLCVHSASPEDV 360
QY 361 ALYVGLIAVAVCLVLLVLLVLYCRKKEGLSDVADSSILTSFGFQVSIKPSKADNPHL 420
Db 360 ALYVGLIAVAVCLVLLVLLVLYCRKKEGLSDVADSSILTSFGFQVSIKPSKADNPHL 419
QY 421 LTIQPDLS-TTTTYQGSCLPRQDGPSPKFQLTNGHLLSPGGRHRLHSSPTSEAEFV 479

```
Db 420 LTIQPDLSLTITTYQGSLSRQDGPSPKQFQLSNGHLLSPGSGRHTLHSSPTSEAE 479
QY 480 SFLSTQNYFRSLPRGTSNMTYGTNFGLGRMLPNTGISLLIPPDAPRGKIYEIYLT 539
Db 480 SFLSTQNYFRSLPRGTSNMTYGTNFGLGRMLPNTGISLLIPPDAPRGKIYEIYLT 539
QY 540 KPEDVRLPLAGCQTLSPVSCGPPGVLTRPVILAMDHCGEPPDSWSLRKKQSCGS 599
Db 540 KPEDVRLPLAGCQTLSPVSCGPPGVLTRPVILAMDHCGEPPDSWSLRKKQSCGS 599
QY 600 WQDVLHLGEEAPSHLYYCOLEASACVFTQELGRFALVGEALSVAAAKRLKLLFAPVA 659
Db 600 WE-DVLHLGEEAPSHLYYCOLEAGACVFTQELGRFALVGEALSVAAAKRLKLLFAPVA 658
QY 660 CTSLEYNIRVYCLHDTHDALKEVVQLEKQGGQLIQEPVLFHFKDSYHNLRLSHDVPSS 719
Db 659 CTSLEYNIRVYCLHDTHDALKEVVQLEKQGGQLIQEPVLFHFKDSYHNLRLSHDVPSS 718
QY 720 LWSKLLVSYQEIPIFFHYHNGTQRYLHCTFTLERSVSPSTDACKLWVQVGDGQSPSI 779
Db 719 LWSKLLVSYQEIPIFFHYHNGTQRYLHCTFTLERSVSPSTDACKLWVQVGDGQSPSI 778
QY 780 NNINIKDTRFAELLALAESEAGVPALVGPSPAFKIPFLIRQKIISLDPCCRRGADWRTLAQ 839
Db 779 NNINIKDTRFAELLALAESEAGVPALVGPSPAFKIPFLIRQKIISLDPCCRRGADWRTLAQ 838
QY 840 KLHLDLSLSPFASKPSPTAMILNLWEARHPFNGNLSQAAAAVAGLQPDAGLFTVSEARC 899
Db 839 KLHLDLSLSPFASKPSPTAMILNLWEARHPFNGNLSQAAAAVAGLQPDAGLFTVSEARC 898

RESULT 3
Q96GP4 PRELIMINARY; PRT; 544 AA.
AC Q96GP4;
DT 01-DEC-2001 (TremBLrel. 19, Created)
DT 01-DEC-2001 (TremBLrel. 19, Last sequence update)
DT 01-JUN-2003 (TremBLrel. 24, Last annotation update)
DE Similar to transmembrane receptor Unc5H1 (Fragment).
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1]_
RP SEQUENCE FROM N.A.
RC TISSUE=Brain;
RA Strausberg R.;
RL Submitted (JUN-2001) to the EMBL/GenBank/DDBJ databases.
DR EMBL; BC009333; AA09333.1; -.
DR GO; GO:0016021; C:integral to membrane; IEA.
DR GO; GO:0004872; F:receptor activity; IEA.
DR GO; GO:0007165; P:signal transduction; IEA.
DR InterPro; IPR000488; Death.
DR InterPro; IPR000906; ZUS.
DR Pfam; PF00531; death; 1.
DR Pfam; PF00791; ZUS; 1.
DR SMART; SM00005; DEATH; 1.
DR SMART; SM00218; ZUS; 1.
KW Receptor; Transmembrane.
FT NON TER 1
SQ SEQUENCE 544 AA; 59949 MW; 350A7BA53373CCAE CRC64;

Query Match 58.9%; Score 2819; DB 4; Length 544;
Best Local Similarity 99.6%; Pred. No. 8.2e-250;
Matches 540; Conservative 0; Mismatches 2; Gaps 2;

QY 359 DVALVGLIAVAVCLVLLLVILVYCRKKGLEDSDVADSSILTSFGQVSVTKPSKADNP 418
Db 4 DVALVGLIAVAVCLVLLLVILVYCRKKGLEDSDVADSSILTSFGQVSVTKPSKADNP 63
QY 419 HLLTIQPDLS-TTTTYQGSLSRQDGPSPKQFQLTNGHLLSPGSGRHTLHSSPTSEAE 477
```

```
Db 64 HLLTIQPDLSLTITTYQGSLSRQDGPSPKQFQLTNGHLLSPGSGRHTLHSSPTSEAE 123
QY 478 FVSRSTQNYFRSLPRGTSNMTYGTNFGLGRMLPNTGISLLIPPDAPRGKIYEIYLT 537
Db 124 FVSRSTQNYFRSLPRGTSNMTYGTNFGLGRMLPNTGISLLIPPDAPRGKIYEIYLT 183
QY 538 LHKPEDVRLPLAGCQTLSPVSCGPPGVLTRPVILAMDHCGEPPDSWSLRKKQSCGS 597
Db 184 LHKPEDVRLPLAGCQTLSPVSCGPPGVLTRPVILAMDHCGEPPDSWSLRKKQSCGS 243
QY 598 GSWEQDVLHLGEEAPSHLYYCOLEASACVFTQELGRFALVGEALSVAAAKRLKLLFAP 657
Db 244 GSWE-DVLHLGEEAPSHLYYCOLEASACVFTQELGRFALVGEALSVAAAKRLKLLFAP 302
QY 658 VACTSLEYNIRVYCLHDTHDALKEVVQLEKQGGQLIQEPVLFHFKDSYHNLRLSHDVP 717
Db 303 VACTSLEYNIRVYCLHDTHDALKEVVQLEKQGGQLIQEPVLFHFKDSYHNLRLSHDVP 362
QY 718 SLSKLLVSYQEIPIFFHYHNGTQRYLHCTFTLERSVSPSTDACKLWVQVGDGQSPSI 777
Db 363 SLSKLLVSYQEIPIFFHYHNGTQRYLHCTFTLERSVSPSTDACKLWVQVGDGQSPSI 422
QY 778 SINFINIKDTRFAELLALAESEAGVPALVGPSPAFKIPFLIRQKIISLDPCCRRGADWRTLAQ 837
Db 423 SINFINIKDTRFAELLALAESEAGVPALVGPSPAFKIPFLIRQKIISLDPCCRRGADWRTLAQ 482
QY 838 AOKLHLDLSLSPFASKPSPTAMILNLWEARHPFNGNLSQAAAAVAGLQPDAGLFTVSEARC 897
Db 483 AOKLHLDLSLSPFASKPSPTAMILNLWEARHPFNGNLSQAAAAVAGLQPDAGLFTVSEARC 542
QY 898 EC 899
Db 543 EC 544

RESULT 4
O08747 PRELIMINARY; PRT; 931 AA.
AC O08747;
DT 01-JUL-1997 (TremBLrel. 04, Created)
DT 01-JUL-1997 (TremBLrel. 04, Last sequence update)
DT 01-OCT-2003 (TremBLrel. 25, Last annotation update)
DE Rostral cerebellar malformation protein.
GN UNCSH3 OR RCM.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
RN [1]_
RP SEQUENCE FROM N.A.
RC STRAIN=C57B6/SJL;
RX MEDLINE=97271898; PubMed=9126743;
RA Ackerman S.L., Kozak L.P., Przyborski S.A., Rund L.A., Boyer B.B.,
RA Knowles B.B.;
RT "The mouse rostral cerebellar malformation gene encodes an UNC-5-like
RT protein.";
RL Nature 386:838-842(1997).
DR EMBL; U72634; AAB54103.1; -.
DR MGD; MGI:1095412; Unc5h3.
DR GO; GO:0005886; C:plasma membrane; IC.
DR GO; GO:0005042; F:netrin receptor activity; IDA.
DR GO; GO:0005515; F:protein binding; IDA.
DR GO; GO:0007420; P:brain development; IMP.
DR GO; GO:0030334; P:regulation of cell migration; IMP.
DR InterPro; IPR000488; Death.
DR InterPro; IPR007110; Ig-like.
DR InterPro; IPR003598; Ig_c2.
DR InterPro; IPR008884; TSP1.
DR InterPro; IPR008085; TSP1.
DR InterPro; IPR000906; ZUS.
DR Pfam; PF00531; death; 1.
DR Pfam; PF00047; ig; 1.
DR Pfam; PF00090; tsp_1; 2.
```


Qy	186	NEDLVDP	SLDPNVYITRBSHVVROARLADTANYTCVAKNIIVARRRSAAAIIVYNGW	245
Db	204	NEDIIDP	ADRNFYITI DHNLIIKQARUSDYTANYTCVAKNIIVAKRSTTATVIYVNGW	263
Qy	246	STWTWSVCSAS	CGRGWOKRSRSCNTPAPLNGGAFCEGONVHDRTVSSLLVSVDGSWPW	305
Db	264	STWTWSVCNSRCRGYQKRTRCTNPAPLNGGAFCEGQS-V	KIACTTLCVPDGEFTSW	322
Qy	306	SKWASGLDCTHWRSRECDSPAPRNGGECQTGDTDLTRNCTSDLCVH		352
Db	323	SKNSTCGTECHWRRRRECTAPAPKNGKDCDGLVLQSKNCTDGLCMQGFIPISTEHRQP		382
Qy	353	-----SASGPEDVALYVGL-IAVAVCLVLLLILVLYCRKKEGLSDVADSSILTSGF		405
Db	383	NEYGFSASPDDVALYGVIVTAVTCLAITVVALFYVKXNRFPESDIIDSSALNGF		442
Qy	406	QPVSIKPSKANPHLLTQPD-LSTTTTYQGLSCPQDGSPKFOLTGNHLSPGLGGRH		464
Db	443	QPVNIKAAQQD---LLAVPPDLTSAAMVRGVPVYALHD-VSDKIPTMNSPIIDLPLNLKI		498
Qy	465	TLEHSS----PTSEAEFEVSRLS---TONYF-----RSLPRGT--SNMTYCTFPNEL		506
Db	499	KVYNSSGAVTQDDLAEFSSKLSPOMTSLEENEAINLKQLARQTDPCSTAFGTFNSL		558
Qy	507	GGRIMPTNGISLLIPPDPAIPRGKIYEIYLTLHKPEDVRLPLAGQOTLLSPIVSCGPPGV		566
Db	559	GGHLIIPNSGVSLIIPAGAIPOGRVYEMVTVHRKENMRPMEDSQTLITPVVSCGPGA		618
Qy	567	LLTRPVILAMDHCCPEPDSNSLRLLKKQSCGSEWDVLHGEEAPSHLIYYCOLEASAC		626
Db	619	LLTRPVILTLHCADPSTEDWKIQIKNAQVQOWE-DVVVGVEENFTTFCYIQLDAEACH		677
Qy	627	VFTFQLGRFALVGBALSVAATAKRLKLLFPVACTSLEYNIKVYCLDTHDKAEVVOLE		686
Db	678	ILTENISTYALVGQSTTKAAARKUKLAIFGPLCCSSLEYSIRVYCLDQDALKEVUQE		737
Qy	687	KQLGGOLIQEPRVLHFQDSYINLRLSIHDFVSSLWKSLLVSYQBIFPYHIWNGTQRYLH		746
Db	738	RQMGGULLEEPKALRFKGISIHNLRLSIHDIASHLSWKSLLAKYQBIFFVHIWSGQRNLH		797
Qy	747	CTFTLERVSPSTDACKLWWQEGDGSFSINFENITKDTEFAELLASEAGYPALVG		806
Db	798	CTFTLERLSNTVELVCKLVCPQVEGEGQIFQLNCTVSEEPGTGIDPLDLPASTITVTG		857
Qy	807	PSAFPIPLEIROKIISSLDPPCRRGADWRTLAQKLHLDShLSFFASKPSPTAMILNLWEA		866
Db	858	PSAFSIPLEPIROKLCSSLDAPOTRGHDWELAHKLNLDRYLNYFATKSPSGTGIVILLWEA		917
Qy	867	RHPFNIGNLSQAAAAVAGLGQPDAGLFTVSEAE	898	
Db	918	QNFPDGNLSMLAAVLEEMGRHETVTVYLAAGQ	949	
RESULT 6				
Q7T2Z5	ID	Q7T2Z5	PRELIMINARY;	PRT; 931 AA.
AC	Q7T2Z5			
DT	01-OCT-2003	(TREMBLrel. 25, Created)		
DT	01-OCT-2003	(TREMBLrel. 25, Last sequence update)		
DT	01-OCT-2003	(TREMBLrel. 25, Last annotation update)		
OS	UNC5-like protein 3.			
OS	Gallus gallus (Chicken).			
OC	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;			
OC	Archosauria; Aves; Neognathae; Galliformes; Phasianidae; Phasianinae;			
OC	Gallus.			
NCBI_TaxID=9031;				
[1]				
SEQUENCE FROM N.A.				
RA	Guan W., Condic M.L.;			
PT	"Characterization of Netrin-1, Neogenin and cUNC-5H3 expression during			
RT	chick dorsal root ganglia development."			
RL	Gene Expr. Patterns 3:369-373(2003).			

DR	ENBL	AY187310	AA067275.1	-	
SO	SEQUENCE	931 AA	102906 MW	1E23A0D84F2E2C62	CRC64
	Query Match	56.0%	Score 2679.5	DB 13	Length 931
	Best Local Similarity	56.1%	Pred. No. 1.6e-238		
	Matches	511	Conservative	151	Mismatches 218
				Indels	31
				Gaps	12
QY	9	PALLGIVLA	AWLRGSGAQQS	---	ATVANPVGANPDILLPHFLVPEDVYIVKKNPVLVLC 65
DB	26	PAL	-AVLGASRFGSAAQDDFFHELPETFPSPDPPEPLPHFLIEPEEAYIVKKNPVLVLC 83		
QY	66	KAVPATQIFK	CNCEWYRQVDHVIERTDSSGSEPTMEVRINTVSRQVKEVGLVEEWQC 125		
DB	84	KASPATQIY	FKCNSEWYQKHVVDERVDTSLGIVCEVSEIERSQQVESELPFGEDWQC 143		
QY	126	CVAWSSGTTKS	QKAYIRIARLRNPFQEPLEAKESVLEQGVILFCRPEGGIPPAEVEWLR 185		
DB	144	CVAWSSAGTTK	SRKAYVRIAYLRKTFEQEPLEKESVLEQEVLLQCRPPEGIPPAEVEWLR 203		
QY	186	NEDLVPSLD	PNVYITFEHSLVVRQARLADTANYTCVAKNIVARRSASAAVIVYNGW 245		
DB	204	NEEDVPE	VDNRNYITIDHLLIKQRLSTANYTCVAKNIVARRSSTTATVIVYNGW 263		
QY	246	STWTWSVC	SASGCRGQWKEKRSRCTNPAPLNGGAFCEGQNVHRTVSSLVSVVDGMSWP 305		
DB	264	STWTWSAC	NRGCRGFQKTRCTNPAPLNGGAFCEGQNV-QKIACITTLCPVDGKWTW 322		
QY	306	SKNSACGLD	CTHWRSECDPAPNGBEGCQGTDLTRNCTSLCVHSASGPEDVALYVG 365		
DB	323	SKWSTCGTE	CHWRRECTAPAPKNGKDCGEVLQSKNCTDGLCMQAAPSDDDVALYVG 382		
QY	366	L-TAVAVCL	VLLVLLVYVCRKKEGLSDVDASSILTSGFQPVSVIKPSKADNPILLTIQ 424		
DB	383	IVTAVIVCL	AISSVVALFVYRKNHRFPESDIUSSALNGGFQPVNIAKARQD---LNAV 439		
QY	425	PDL-STTTTY	QGSCLPRQDGPSPKFQTLNGHLLSLPLGGRRHTLHSS---PTSEAEFV 479		
DB	440	PDLTSAAM	VRGPVYALHD-VSDKIPMTNSPILDPPLNLKIKVYNTSGAVTPQDELSDFS 498		
QY	480	SRLS---	TONYF-----RSLPRGT--SNMYGTGFNPLGRLMPTNGISLLIPDDA 525		
DB	499	SKUSPQIT	QSLLENETINVNKQSLARQTDPSCTAFGTNSLGHVLVPNSGVSLIPAGA 558		
QY	526	IPRGKTYE	ILTLHKPEDVRLPLAGQOTLLSPVSCGPPGVLLTRPVLAMDHCGEPPD 585		
DB	559	VPQGRVY	VMVTVHRKEGMRPPVEDSQTLTPVVS CGPFGALLTRPVLTMHCAEPNMD 618		
QY	586	SWSLRLK	KQSCGSDQDVLHGEAPSHIYQCLREASACYVTEQGRPALVGEALSV 645		
DB	619	DWQIQIKH	QAGQGFNB-DVVVGEENFTTFCYIQLDPEACHILTELTSTVALYVQSITKA 677		
QY	646	AAKRLKLL	FAPVACTSLEYNIRVYCLHDTHDALKEVVOLKQGLQIQEPRVLHFKDS 705		
DB	678	AAKRLKAI	FGLPSCSSELYSIEVYCLDQTDALKEVQLERQMGQQLLEEPKILHFKGS 737		
QY	706	YHNLRLS	IHDVPESSLWKSLLVSYQBIPIFYHWNQYRLHCTFTLERSVSPSTDIAKCL 765		
DB	738	THNLRLS	IHDIAHSLWKSLLPAKYQBIPIFYHWSGQRNLHCTFTLERTSLNLTLEVCVKL 797		
QY	766	WYVQVGE	CGQSFINNITKDTFFAELLALLESAGYPALVGSAPKIPFLIRQKILSSLD 825		
DB	798	CVRQVE	GEQCIQIQLNCSVEETGIDPYIMDSAGSITTIIVGNAFSIPILPIRQKCSSLD 857		
QY	826	PPCRRG	ADWRTIAQKILHLSLHFFSFKSPSTAMILNLWEARHPFNNGNISQLAAAAVAGLG 885		
DB	858	APQTRGD	WRMLAHKILDRYLNIFATKSSPTGVILLDWEAQNFDPGNLSMLAAVLEEMG 917		
QY	886	QPDAGL	FTVSE 896		
DB	918	RHETVV	SLAAE 928		

```
O95185
ID O95185 PRELIMINARY; PRT; 931 AA.
AC O95185;
DT 01-MAY-1999 (TrEMBLrel. 10, Created)
DT 01-MAY-1999 (TrEMBLrel. 10, Last sequence update)
DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
DE Transmembrane receptor UNC5C.
GN UNC5C.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1]_
RP SEQUENCE FROM N.A.
RC TISSUE=Brain;
RX MEDLINE=99000841; PubMed=9782087;
RA Ackerman S.L., Knowles B.B.;
RT "Cloning and mapping of the UNC5C gene to human chromosome 4q21-q23.";
RL Genomics 52:205-208(1998).
DR EMBL; AF055634; AAC67491.1; -.
DR Genew; HGNC:12569; UNC5C.
DR GO; GO:0005042; F-actin receptor activity; TAS.
DR GO; GO:0007411; P-axon guidance; TAS.
DR GO; GO:0007420; P-brain development; TAS.
DR InterPro; IPR000488; Death.
DR InterPro; IPR007110; Ig-like.
DR InterPro; IPR003598; Ig_c2.
DR InterPro; IPR000884; TSP1.
DR InterPro; IPR000885; TSP1.
DR InterPro; IPR000906; ZU5.
DR Pfam; PF00531; death; 1.
DR Pfam; PF00047; ig; 1.
DR Pfam; PF00090; tsp_1; 2.
DR Pfam; PF00791; ZU5; 1.
DR PRINTS; PR01705; TSP1REPEAT.
DR SMART; SM00005; DEATH; 1.
DR SMART; SM00408; IGC2; 1.
DR SMART; SM00209; TSP1; 2.
DR SMART; SM00218; ZU5; 1.
DR PROSITE; PS00835; IG-LIKE; 1.
DR PROSITE; PS0092; TSP1; 2.
KW Immunoglobulin domain; Receptor.
SQ SEQUENCE 931 AA; 103101 MW; EFD71122C98DAB8 CRC64;

Query Match 55.8%; Score 2673.5; DB 4; Length 931;
Best Local Similarity 55.8%; Pred. No. 5.7e-238;
Matches 507; Conservative 154; Mismatches 221; Indels 31; Gaps 12;

QY 9 PALLGIVLAALRGSGAQQS---ATVANPVPGANPDLLPHFLVEPEDVIVKPKVLLVC 65
DB 26 PAL--ALLSASGTGSAQDDPFHELPETFPPDPPEPLPHFLIEPEAVIKKPKVNLVC 83
QY 66 KAVPATQIPKCKNGEWWRVQDVHIERSTGSSGEPTWMEVRINVSROQVEKVFGLBEYWCQ 125
DB 84 KASPATQIQPKCKNGEWWRVQDVHIERSTGSSGEPTWMEVRINVSROQVEKVFGLBEYWCQ 143
QY 126 CVAWSSGTTKSKAYIRIARLKNFEQELAKEVLSLEQIVLPCRPPEGIPPAEVEWLRL 195
DB 144 CVAWSSAGTTKSKAYIRIARLKNFEQELAKEVLSLEQIVLPCRPPEGIPPAEVEWLRL 203
QY 186 NEDLVDPSPDPNVIYITREHSLVVRQARLADTANYTCVAKNIIVARRRSASAAVIVYNGGW 245
DB 204 NEDIIDPVEDRNFYITIDHNLIIKQARLSDTANYTCVAKNIIVARRRSASAAVIVYNGGW 263
QY 246 STTWTSVCSASCGRGKRSRSTNPAFLNGAFCEGQNVHRTVSSLLSVSDGWSWPW 305
DB 264 STTWTSVCSNCRGCGYKQKTRCTNPAFLNGAFCEGQSV-QKIACTTLCPVDGRWTPW 322
QY 306 SKWSACGLDCTHWRSCSPAPRNGECGQTDLTNCTSLCVHSASGCPEDVALYVG 365
DB 323 SKWSTCTECHWRRECTAPAPKNGKCDGLVLSQKNCITDGLCQWATPDSDDVALYVG 382
QY 366 L-IAVAVCLVLLLVLLVILVYCRKEGLSDVADSSILTSGFQPVSIKPSKADNPHLLTIQ 424
```

```
Db 383 IIVAVIVCLAISVVVALFVYRKNHRDFESDIIDSSALNGGFQPVNKAARQD---LLAVP 439
QY 425 PDL-STTTTQGSCLCPRGDPSKFLTNHLLSLPLGGRRHLLHS-----SPTSAREEV 479
Db 440 PDLTSAAMYRGVPVALHD-VSDKIPMTNSPILDPNLKIKYVNTSGAVSQDDLSERT 498
QY 480 SRLS---TQNYF-----RSLPRGT--SNMTYGTENFLGRLMIPMTGISLLIPDDA 525
Db 499 SKLSQMTQSLLENALSILKNQSLARQDPTSCATGFSNLSGLHLLVFNPSGVSLIPAGA 558
QY 526 IPRGIYEIYLTLHKPEDVRLPLAGCQTLLSPVSGPPEGVLLTRPVILAMDHQGPSPD 585
Db 559 IPQGRVYEMYVTVHRKETMRPPMDSDQTLLTPVWSCGPGALLTRPVVLTMRHCADPTE 618
QY 586 SWSLRLKQSCGSEWQDVHLHGEAPSHLYLCOLEASACYVFTGOLGRFALVGEALSVA 645
Db 619 DKWILLKQAQAGQME-DVVVVGEEENFTPCYIKLDAEACHILLENLSYALVGHSTTKA 677
QY 646 AAKRLKLLFAPVACTSLEYNIRVYCLDTHDALKEVQLEKQLGGQLIQEPRVLHFKDS 705
Db 678 AAKRLKLAIFGLPCCSSLEYSIRVYCLDQTQALKELHLERQTGGQLLEPKALHFKGS 737
QY 706 YNLRLSIHDVPSSLWKSLLVSYQEIIPYHIIWNGTQRYLHCTFTLERSVSPSTSLACKL 765
Db 738 THNLRLSIHDIAHSLWKSLLAKYQEIIPYHVMWSGQRNLHCTFTLERSVSLNTVELVCKL 797
QY 766 WYVWEGDQGSFISFNITKTRFAELLALASEAGVPALVGPSAFKIPFLIKLISSLD 825
Db 798 CVRQVEGEGQIFQLNCTVSEETGIDPLDLPANTITTVTGPSAFSIPFIRKQKCSSLD 857
QY 826 PCRRGADWRTLAQKLHLDLSHLSFFASKPSPTAMILNLWEARHPFNGLNSQLAAAVAGLG 885
Db 858 AFQTRGHDWRMLAHKLNLDRLNYFATKSSPTGVILDLEAQNFPDGNLSMLAAVLEENG 917
QY 886 QPDAGLFTVSEAE 898
Db 918 RHETVVSAAEGQ 930

RESULT 8
Q8JGT4 PRELIMINARY; PRT; 943 AA.
AC Q8JGT4;
DT 01-OCT-2002 (TrEMBLrel. 22, Created)
DT 01-OCT-2002 (TrEMBLrel. 22, Last sequence update)
DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
DE UNC-5 receptor.
OS Xenopus laevis (African clawed frog).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Amphibia; Batrachia; Anura; Mesobatrachia; Pipidae;
OC Xenopodinae; Xenopus.
OX NCBI_TaxID=8355;
RN [1]_
RP SEQUENCE FROM N.A.
RA Anderson R.B., Holt C.E.;
RT "Expression of UNC-5 in the developing Xenopus visual system.";
RL Submitted (APR-2002) to the EMBL/GenBank/DBJ databases.
DR EMBL; AY09459; AAM34486.1; -.
DR GO; GO:0004872; F-actin receptor activity; IEA.
DR GO; GO:0007165; P-signal transduction; IEA.
DR InterPro; IPR003599; Death.
DR InterPro; IPR003599; Ig.
DR InterPro; IPR007110; Ig-like.
DR InterPro; IPR003598; Ig_c2.
DR InterPro; IPR000884; TSP1.
DR InterPro; IPR000885; TSP1.
DR InterPro; IPR000906; ZU5.
DR Pfam; PF00531; death; 1.
DR Pfam; PF00047; ig; 1.
DR Pfam; PF00090; tsp_1; 2.
DR Pfam; PF00791; ZU5; 1.
DR PRINTS; PR01705; TSP1REPEAT.
DR
```



```
Query Match 52.3%; Score 2505.5; DB 11; Length 1008;
Best Local Similarity 52.7%; Pred. No. 2.3e-222;
Matches 497; Conservative 153; Mismatches 240; Indels 53; Gaps 17;

QY 1 MAVRPGMLVALLGIVLAAM-----LRG--SGAQQSATVANPVPGANPDLLPHFLVPEDEV 53
DB 75 MRARSGVRSALLLALLCWDPTPSLAGVDSAGQ---VLPDSYPSAPAEQLPYFLLEPQDA 131
QY 54 YIVKNKPVLLVCKAVPATQIFPKNGEWYRQVDHVIERSTGSSGEPTWVEVNRQOV 113
DB 132 YIVKNKPVLLVCKAVPATQIFPKNGEWYRQVDHVIERSTGSSGEPTWVEVNRQOV 191
QY 114 EKVPGLVEEYWCQVAMSSGTTKSKAYIRIARLKNFQEPDLAKEVSLQGIPLPCRP 173
DB 192 EGVFAEVEWLNKEDVIDPAQDNFLTIIDHNLIRQARLSDTANTYCVAKNIVARRSA 251
QY 174 EGIPPAEVEWLNKEDVIDPAQDNFLTIIDHNLIRQARLSDTANTYCVAKNIVARRSA 233
DB 252 EGVFAEVEWLNKEDVIDPAQDNFLTIIDHNLIRQARLSDTANTYCVAKNIVARRSA 311
QY 234 SAAVIVYNGWSTWTEWVSCASCGRWKRSCTNPAPLNGAFCEGQVHRTVSS 293
DB 312 TATVIVYNGWSSWAENPFCNRCGRGWKRTCTNPAPLNGAFCEGQ-AFQKTACT 370
QY 294 LLVSDVGSWSPKWSACGLDCTHWRSCSDPAPNGGEBECQGTDLTRNCTSDLCV 353
DB 371 TVCPVDGAWTEWKSACSTECARWRSRECMAPPQNGGRDCSGTLLDKNCTDGLCVLN 430
QY 354 ASGPEDVALYVGL-TAVAVCLVLLLVILVYCRKKEGLSDVDSS--ILTSGFQPVSTK 411
DB 431 LETSGDVALYAGLVAVFVVAVLMAVGIVVYRNRCDFTDITDSSAALTGFHFVNEK 490
QY 412 PSKADPHLL--TIQDLSLTTT--TYOGLCPRODGPSPKQFQTLNGLHLLSPGGRTHLH 468
DB 491 TARPNNPQLLHPSAPDLTASAGIYRGPVYALQDS-ADKIPMTNSPLDPLPSLKIKVTN 549
QY 469 SS-----PTSEAEFVSRSLSTQNYFRS-----LPRGTSNMT 499
DB 550 SSTIGSGGLADGADLLGVLPQTYPGDP-SRDTHFLHRSASLGSHLLGLPRDPSSV 608
QY 500 YGTFFNLGGLRMLPNTGISLIPDAIPRGKIYEIYTLHKPEDVRLPLA-GQOTLLSPI 558
DB 609 SGTFGLGGLRLSLPGTGVSLVFNPAIPQKPFYDLVHINKABST-LPLSEGSQTVLSPS 667
QY 559 VSCGPGVLLTRVILAMDHCGPSPDSLSRLKQSCGSEWQDVHLHGEERAPSHLYC 618
DB 668 VTCGPTGLLLCRPVLLTVPHCAEVIAGDMTIFOLKTAHQHWP-EVVTLDDETLNTPCYC 726
QY 619 QLEASACYFTEGLGRFALVGEALSVAAAKRLKLLPAPVACTSLEYNIRVYCLHDTDA 678
DB 727 QLEAKSCHILLDQGLTYVFMGESYSRSVAVRLQAIAPALCTSLYSRLVYCLEDTFVA 786
QY 679 LKEVVOLEKQGLQIQEPRVLHFQDSYHNRLSIHDVPSSLWKSLLYSQYIPIFYHIW 738
DB 787 LKEVLELERTLGGVLYVEEPKPLLFKDSYHNRLSLHDIPHAWKSLKAKYQIPIFYHV 846
QY 739 NGTQVILHCTFLERVSPSTSDACKLWVQVGDGQSFNINITYKDRFALLALESE 798
DB 847 NGSQRALHCTFLERHSLASTETFCVCRVQVGEQGIQFQHTTLA-ETPAGSLDLCSEA 905
QY 799 AG--VPALVGPSAFKIPFLRQKIISLDPDPCRGADWRTLAQKLHLDHLSLFFASKPSP 856
DB 906 PGNAITQLGPAFKIPLSTRQKICSLDAPNRSRGNDWRLLAQKLSMDRYLNYFAIKASP 965
QY 857 TAMILNLWEARHPFNGNLQAAVAGLQDPDAGLFTVSEAC 899
DB 966 TGVILDLWEARQQDDGLNSLASALEMGKSEMLVAMATDGC 1008

RESULT 10
Q8KLS3
ID Q8K183
AC Q8K183;
PRELIMINARY; PRT; 945 AA.
```

```

QY 401 LTSGFQPVSIKPSKADNPHELL--TIQPDLSITT--TYQGLSCPQDGPSPKFKQLTNGHLLS 457
Db 417 LTGGFHPVNFKARPNNPQLLHSPADPLTASAGIYRGVPYALQDS--ADKIPMTNSPLLD 475
QY 458 PLGGGRHTLHSS-----PTSEAEFVSRLSTQNYFRS----- 490
Db 476 PLPSLKIKVYNSSTIGSGGLADGADLLGLVPGTYPGDF--SRDTHFLHLSASLSQHL 534
QY 491 --LPRGTSNMTYGTFFLAGELMIPNTG:SLIPDPAIRPKIYEIYLILHFKPEDVRLPL 548
Db 535 LGLPRDPSSVSTGFCGLGRSLPGLTASAGIYRGVPYALQDS--ADKIPMTNSPLLD 593
QY 549 A--GCOTLLSPIVSCGPPGVLTPVILAMDHCGEPPSDSLSLKKQSCGSGWEQDVLHL 607
Db 594 SEGSQTVLSFVTCGTGTLCPVVLTPVHCAEVIAGDWIQLKTAHQHWE--EVTIL 652
QY 608 GBEAPSHLYYCOLEASACVYFTQGRFALVGEALSAAAKRLKLLFAPVACTSLEYNI 667
Db 653 DEETLNTPCYCOLEAKSCHILLDQLGTYVFMGESYSRSYSAVKRLQLAIFAPALCTSLEYS 712
QY 668 RVYCLDTHDALKEVVQLEKQGLQIOPRVILHFKDSYHNLRLSHDVPSSLWKSLLV 727
Db 713 RVYCLEDTVALKEVLELERTLGGYLVBEKPLLFKDSYHNLRLSHDIPHAHRSKLLA 772
QY 728 SYQEIPFYHIWNGTORYLHCTFTLERSVSPSTSDLACKLWVWQVGGQSGFSINFINIKDT 787
Db 773 KYQEIPFYHVMNGSQRALHCTFTLERSVSPSTSDLACKLWVWQVGGQSGFSINFINIKDT 831
QY 788 RFAELLALSEAG--VPALVGSFAKIFPLIRQKIISLSDPPCRGADWRTLAQKLHLS 845
Db 832 PAGESDALCSAPCNAITTLQGLYAFKIPLSIRQKICSSLDAPNSRGNDWLLAQLKSMR 891
QY 846 HUSFEASKPSPTAMILNLWEARHFNGLNLSQLAAGVAGLQPDAGLFTVSEAC 899
Db 892 YLNYFATKASPTGVILLDWEARQQDGLNLSALAEEMGKSEMLVAMATDGC 945

RESULT 11
O08722 PRELIMINARY; PRT; 945 AA.
AC O08722;
DT 01-JUL-1997 (TREMBLrel. 04, Created)
DT 01-JUL-1997 (TREMBLrel. 04, Last sequence update)
DT 01-OCT-2003 (TREMBLrel. 25, Last annotation update)
DE Transmembrane receptor UNC5H2.
OS Rattus norvegicus (Rat).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.
OX NCBI_TaxID=10116;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=97271897; PubMed=9126742;
RA Leonardo E.D., Hnack L., Masu M., Keino-Masu K., Ackerman S.L., Tessier-Lavigne M.;
RT "Vertebrate homologues of C. elegans UNC-5 are candidate netrin receptors.";
RL Nature 386:833-838 (1997).
DR EMBL: U87306; AAB57679.1; -.
DR GO: GO:0004872; F:receptor activity; IEA.
DR GO: GO:0007165; P:signal transduction; IEA.
DR InterPro: IPR000488; Death.
DR InterPro: IPR007110; Ig-like.
DR InterPro: IPR003598; Ig C2.
DR InterPro: IPR000884; TSP1.
DR InterPro: IPR000805; TSP1.
DR InterPro: IPR000906; ZUS.
DR Pfam: PF00531; death; 1.
DR Pfam: PF00047; ig; 1.
DR Pfam: PF00090; tsp_1; 2.
DR Pfam: PF00791; ZUS; 1.
DR PRINTS: PR01705; TSP1REPEAT.
DR SMART; SM00005; DEATH; 1.

```

```

DR SMART; SM00408; IGC2; 1.
DR SMART; SM00209; TSP1; 2.
DR SMART; SM00218; ZUS; 1.
DR PROSITE; PS0017; DEATH DOMAIN; 1.
DR PROSITE; PS00835; IG LIKE; 1.
DR PROSITE; PS0092; TSP1; 2.
KW Immunoglobulin domain; Receptor.
SQ SEQUENCE 945 AA; 103520 NW; 659C2A262E560B9B CRC64;

Query Match 52.2%; Score 2497; DB 11; Length 945;
Best Local Similarity 52.0%; Pred. No. 1.3e-221;
Matches 501; Conservative 144; Mismatches 236; Indels 82; Gaps 20;

QY 1 MAVRPLMALLIGVLAAM-----LRG--SGAQQSATVANVPVGNPDLLPHFLVPEDV 53
Db 1 MRARSARGALLALLLCWDPTSLAGIDSGGQ--ALPDSFPAPAEQLPHFLLEPEDA 57
QY 54 YIVKPKVLLVCKAVPATQIFFKCNGEYVQVDHVIERTDSSGSEPTMEVRINVSROOV 113
Db 58 YIVKPKVELHCRATPATQIYFKCNGEWSQKHVTQESLDEATGLRIREVQIEVSRQV 117
QY 114 EKVFGLSEYWCQVAMSSSGTTKSKAYIRIARLRNFEQELAKEVSLQGIIVLPCRP 173
Db 118 EELFGLDLYWCQVAMSSSGTTKSRAYIRIARLRNFEQELAKEVSLQGIIVLPCRP 177
QY 174 EGIPPAEWEMLRNEDLVDPDNPVITREHSLVVRQARLADTANTVCVAKNIVARRSA 233
Db 178 EGVPAEWEMLKNEVDIDPAQDTNELLTIHNLRIARLRNFEQELAKEVSLQGIIVLPCRP 237
QY 234 SAAVIVYVNGWSTWTEWSVCASCGRWKRSCTNPAPLNGGAFCEGQNVHDTVSS 293
Db 238 TATVIVYVNGWSSWAEWSPCNRGCGWQKTRCTNPAPLNGGAFCEGQ--ACQKTACT 296
QY 294 LNVSDGWSWMSKWSACGLDCTHWRSECDPAPNGEGECQGTDLDRNCTSDLCV-- 351
Db 297 TVCPVDGAWTEWSKWSACSTECAHWRSECMAPPQNGGRCDSGTLDSKNTDGLCVLN 356
QY 352 --HSASGE-----DVALYVGL--IYAVCLVLLVLLVLYCRKKEGLDSVADSS--I 400
Db 357 QRTLNDPKSRPLEPSGDVALYAGLVAVFVLAIVLAVGVIVYRNCRDFDITDSSAA 416
QY 401 LTSGFQPVSIKPSKADNPHELL--TIQPDLSITT--TYQGLSCPQDGPSPKFKQLTNGHLLS 457
Db 417 LTGGFHPVNFKARPNNPQLLHSPADPLTASAGIYRGVPYALQDS--ADKIPMTNSPLLD 475
QY 458 PL-----GGG-----RHTLHSSPTSEAEFV 479
Db 476 PLPSLKIKVYDSSTIGSGAGLADGADLLGLVPGTYPGDFSRDTHFLHLS----- 526
QY 480 SELSTQNYFRSLPRGTSNMTYGTFFLAGELMIPNTG:SLIPDPAIRPKIYEIYLILH 539
Db 527 ASLGSQ--HLLGLPRDPSSVSTGFCGLGRSLPGLTASAGIYRGVPYALQDS--ADKIPMTNSPLLD 585
QY 540 KPEDVRLPLA--GCOTLLSPIVSCGPPGVLTPVILAMDHCGEPPSDSLSLKKQSCG 598
Db 586 KTEST--LPLSGSQTVLSFVTCGTGTLCPVVLTPVHCAEVIAGDWIQLKTAHQH 644
QY 599 SWEQDVLHGBEAPSHLYYCOLEASACVYFTQGRFALVGEALSAAAKRLKLLFAPV 658
Db 645 HWE--EVTILDEETLNTPCYCOLEAKSCHILLDQLGTYVFTGESYSRSYSAVKRLQLAIFAPA 703
QY 659 ACTSLEYNIYVCLDTHDALKEVVQLEKQGLQIOPRVILHFKDSYHNLRLSHDVP 718
Db 704 LCTSLEYSLRVYCLEDTPAALKEVLELERTLGGYLVBEKPLLFKDSYHNLRLSHDIP 763
QY 719 SLWKSLLVSYQEIPFYHIWNGTORYLHCTFTLERSVSPSTSDLACKLWVWQVGGQSGFS 778
Db 764 AHWRSKLLAKYQEIPFYHVMNGSQRALHCTFTLERSVSPSTSDLACKLWVWQVGGQSGFS 823
QY 779 INFNITKDTREFAELLALSEAGVPAAL--VGPSAFKIPFLIRQKIISLSDPPCRGADWRT 836
Db 824 LHTTLA--ETPAGSLDALCSAPGNAITTLQGLYAFKIPLSIRQKICSSLDAPNSRGNDWRL 882

```

837 LAQKLDHSLSPFRASKPSPTAMILNWEARHPNGNLSQAAAAVAGLQDPDAGLFTYSE 896
 883 LAQKLSMDRYLNYFATKASPTGVLIDWEARQDDGLNLSLASEENGKSEMLVAMTID 942

897 AEC 899
 943 GDC 945

RESULT 12
 Q9D398
 ID Q9D398 PRELIMINARY; PRT; 945 AA.
 AC Q9D398;
 DT 01-JUN-2001 (Tremblrel. 17, Created)
 DT 01-JUN-2001 (Tremblrel. 17, Last sequence update)
 DT 01-OCT-2003 (Tremblrel. 25, Last annotation update)
 DE 6330415E02Rik protein.
 GN UNC5H2 OR 6330415E02RIK.
 OS Mus musculus (Mouse).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
 OX NCBI_TaxID=10090;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=C57BL/6J; TISSUE=Medulla oblongata;
 RX MEDLINE=21085660; PubMed=11217851;
 RA Kawai J., Shinagawa A., Shibata K., Yoshino M., Itoh M., Ishii Y.,
 RA Arakawa T., Hara A., Fukunishi Y., Konno H., Adachi J., Fukuda S.,
 RA Aizawa K., Izawa M., Nishi K., Kiyosawa H., Kondo S., Yamana K. I.,
 RA Saito T., Okazaki Y., Gofobori T., Bono H., Kasukawa T., Saito R.,
 RA Kadota K., Matsuda H. A., Ashburner M., Batalov S., Casavant T.,
 RA Fleischmann W., Gaasterland T., Gissi C., King B., Kochiwa H.,
 RA Kuehl P., Lewis S., Matsuo Y., Nikaido I., Pesole G., Quackenbush J.,
 RA Sakimi L. M., Staubli F., Suzuki R., Tomita M., Wagner L., Washio T.,
 RA Sakai K., Okido T., Furuno M., Aono H., Baldarelli R., Barsh G.,
 RA Blake J., Boffelli D., Bojunga N., Carninci P., de Bonaldo M. F.,
 RA Brownstein M. J., Bult C., Fletcher C., Fujita M., Gariboldi M.,
 RA Gustinich S., Hill D., Hofmann M., Hume D. A., Kamiya M., Lee N. H.,
 RA Lyons P., Marchionni L., Mashima J., Mazzarelli J., Mombaerts P.,
 RA Nordone P., Ring B., Ringwald M., Rodriguez I., Sakamoto N.,
 RA Sasaki H., Sato K., Schoenbach C., Seya T., Shibata Y., Storch K. F.,
 RA Suzuki H., Toyooka K., Wang K. H., Weitz C., Whittaker C., Wilming L.,
 RA Wynshaw-Boris A., Yoshida K., Hasegawa Y., Kawaji H., Kohetsuki S.,
 RA Hayashizaki Y.,
 RT "functional annotation of a full-length mouse cDNA collection."
 RL Nature 409:685-690(2001).
 DR EMBL; AK018177; BAB31108.1; -.
 DR MGD; MGI:894703; Unc5h2.
 DR GO; GO:0007165; P:signal transduction; IEA.
 DR InterPro; IPR000488; Death.
 DR InterPro; IPR007110; Ig-like.
 DR InterPro; IPR003598; Ig_c2.
 DR InterPro; IPR000884; TSPI.
 DR InterPro; IPR008085; TSP 1.
 DR InterPro; IPR000906; ZU5.
 DR Pfam; PF00531; death; 1.
 DR Pfam; PF00047; ig; 1.
 DR Pfam; PF00090; tsp 1; 2.
 DR Pfam; PF00791; ZU5; 1.
 DR PRINTS; PR01705; TSPIREPEAT.
 DR SMART; SM00005; DEATH; 1.
 DR SMART; SM00408; IGC2; 1.
 DR SMART; SM00209; TSP1; 2.
 DR SMART; SM00218; ZU5; 1.
 DR PROSITE; PS50017; DEATH_DOMAIN; 1.
 DR PROSITE; PS50835; IG_LIKE; 1.
 DR PROSITE; PS50092; TSPI; 2.
 KW Immunoglobulin domain.
 SQ SEQUENCE 945 AA; 103725 MW; 43D33B4524E0CBF2 CRC64;

Query Match 52.1%; Score 2493; DB 11; Length 945;
 Best Local Similarity 52.1%; Pred. No. 3e-221;
 Matches 497; Conservative 152; Mismatches 241; Indels 64; Gaps 18;

QY 1 MAVRGLWPALLGIVLAZW-----LRG--SGAQQSATVANPVGANPOLLPHFLVEPEDV 53
 DB 1 MRARGVSALLALLCWDPTSPSLAGVDSAGQ---VLPDSYSPAPAPQLPFELLEPQDA 57
 QY 54 YIVKPKVLLVCVAVPATQIFKFCNGEVRQVDHVIERSTGSGSEPTMEVNRINVSQV 113
 DB 58 YIVKPKVELHCRAPATQIYFKNGEWVSQNDHVTQESLDEATGLRVREQVIEVSQV 117
 QY 114 EKVFGLIEWCQCVAWSSGTTKSKAYIRIARLKNFEOBLAKEVSLQGIIVLPCRP 173
 DB 118 BELFGLIEDYWCQCVAWSSGTTKSRAYIRIARLKNFDEBLAKEVSLQGIIVLPCRP 177
 QY 174 EGIPPAEVEWELNEDLVDPSPNVYITREHSLVVRQARLADTANTVCVAKNIVARRSA 233
 DB 178 EGVPAEVEWELNEDVIDPAQDTNELLIDHNLIIQARLSTANTVCVAKNIVARRST 237
 QY 234 SAAVIVYNGGWSMTWESVCSACSGRWKFSRSTNPAPLNGAFCEGQVHRTVSS 293
 DB 238 AATVIVYNGGWSWAESPCSNRCRGWKRTCTNPAPLNGAFCEGQ-AFQKTACT 296
 QY 294 LLVSDGWSWPKWSACGLDCTHWRSECDPAPRNGEECGTDLDTNCTSDLCV-- 351
 DB 297 TVCPVDGAWTEWSKWSACSTECARHWRSECMAPPQNGRDCSGTLLDKNCTDGLCVLN 356
 QY 352 -----HSASGPEDVALYVGL-IAVAVCLVLLLLLVYCRKKEGLDSDVADSS-I 400
 DB 357 QRTLNPKSHPLETSGDVALYAGLVAVFVVAVLMAEGVIVYRNCRDFDITDSSAA 416
 QY 401 LTSGQPVSIPKSKADNPHLL--TIQPDLSITT--TYQSLCPRQDGPSPKFLTNGHLS 457
 DB 417 LTGGFHPVNFKTRPNPQLHPSAPPDLTASAGIYRGPVYALQDS-ADKIPMTNSPLD 475
 QY 458 PLGGGRHTLHSS-----PTSEAEFFVSLSTQNYFRS----- 490
 DB 476 PLFLSKIKYVNSTTSGSGLADGADLGLVLPPTGPGDF-SRDTHFLHRSASLSQHL 534
 QY 491 --LPRGTSNMTYGTENFLGRLMIPNTGISLIPDAIPRGKIYEIYTLTHKPEDVRLPL 548
 DB 535 LGLPRDPSSVSGTGLGRLSLPTGVSLLVVPNGAIPQGFYDYLHINKAEST-LPL 593
 QY 549 A-GCQTLLSPVSCGPPGVLLTRPVILAMDHCEPSPDSWSRLKKQSCGSEWQDVHL 607
 DB 594 SEGSGTILSPSVTCGPTGLLLCRPVVLTVPHCAEVIAGDWIIFOLKTQAHQHWE-BVVTL 652
 QY 608 GEEASHLYYCOLEASACVFTQOLGRFALVGEALVAAAKELKILLFAPVACTSEYNI 667
 DB 653 DEETLNTPCYCLEAKSCHILLDQLGSYVFMGESYSRSVAKRLQLAIFAPALCTSEYSL 712
 QY 668 RYVCLHDTHDALKEVVQLEKQGLIQEPVILHFKDSYHNLRSLSHDVPSSLWKSLLV 727
 DB 713 RYVCLDTPVALKEVLELERTLGGYLVPEPKPLLFKDSYHNLRSLSHDIPAHWRSKLLA 772
 QY 728 SYQEIPFYHIMGTORYLHCTFTLERSVSPSTDACKLWVQVEGQSGFSINFTKDT 787
 DB 773 KYQEIPFYHVMGSRALHCTFTLERHSLASTEFTCKVCVRQVEGSGQIFQLHTTLA-ET 831
 QY 788 RFAELLALIESEAG--VPALVGPSAFKIPFLIRQKIISLSDPPCCRCGADWETLAKHLDS 845
 DB 832 PAGSLDALCSAPGNATITQGPYAFKIPLSIRKICSSLDADPSRGNDRLLLAQKLSMR 891
 QY 846 HLSFFASKPSPTAMILNWEARHPNGNLSQAAAAVAGLQDPDAGLFTYSEAPC 899
 DB 892 YLNYFATKASPTGVILDLWEARQDDGLNLSLASEENGKSEMLVAMATDGC 945

RESULT 13
 Q8IZJ1
 ID Q8IZJ1 PRELIMINARY; PRT; 934 AA.
 AC Q8IZJ1;
 DT 01-MAR-2003 (Tremblrel. 23, Created)
 DT 01-MAR-2003 (Tremblrel. 23, Last sequence update)
 DT 01-OCT-2003 (Tremblrel. 25, Last annotation update)

DE	Transmembrane receptor UNC5H2.	
OS	Homo sapiens (Human).	
OC	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;	
OC	Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.	
OX	NCBI_TaxID=9606;	
RN	[1]	
RP	SEQUENCE FROM N.A.	
RX	MEDLINE=2246081; PubMed=12359238;	
RA	Komatsuzaki K., Dalvin S., Kinane T.B.;	
RT	"Modulation of G(alpha(2)) signaling by the axonal guidance molecule	
RT	UNC5H2.";	
RL	Biochem. Biophys. Res. Commun. 297:898-905(2002).	
DR	EMBL; AY126437; AA95701.1; -	
DR	GO; GO:0004872; F:receptor activity; IEA.	
DR	GO; GO:0007165; P:signal transduction; IEA.	
DR	InterPro; IPR000488; Death.	
DR	InterPro; IPR003599; Ig.	
DR	InterPro; IPR007110; Ig-like.	
DR	InterPro; IPR003598; Ig c2.	
DR	InterPro; IPR000884; TSP1.	
DR	InterPro; IPR000885; TSP 1.	
DR	InterPro; IPR000906; ZUS_1.	
DR	Pfam; PF00531; death; 1.	
DR	Pfam; PF00047; Ig; 1.	
DR	Pfam; PF00791; ZUS; 1.	
DR	PRINTS; PR01705; TSP1REPEAT.	
DR	SMART; SM00005; DEATH; 1.	
DR	SMART; SM00409; Ig; 2.	
DR	SMART; SM00408; IGC2; 1.	
DR	SMART; SM00209; TSP1; 2.	
DR	PROSITE; PS50835; IG LIKE; 1.	
DR	PROSITE; PS50092; TSP1; 2.	
XW	Receptor.	
SQ	SEQUENCE 934 AA; 102433 MW; 225B3F506D52B780 CRC64;	
	Query Match	
	Best Local Similarity 51.9%; Score 2486.5; DB 4; Length 934;	
	Matches 490; Conservative 149; Mismatches 255; Indels 45; Gaps 16;	
QY	1 MAVRPGLMPLGIVLAAM--LRSGAQQ-SATVANPVPGANPOLLPHFLVPEPDVIV 56	
DB	1 MGARGGARGALLALLCWDPRLSQAGTDSGEVLPDFFPSAPAEPLPYFLOEQDAIV 60	
QY	57 KKKPVLLVCKVAPATQIFKCNQEWVRQVDHVIERSITDGSSEPTMEVRINVSROQVEKV 116	
DB	61 KKKPVELCRAPPATQIYFKCNQEWVSQNDHVTQEGLEATGLRVREVQIEVSRQQVEEL 120	
QY	117 FGLEIYWCQVWSSGTTKQKAVIRIARLKNFEQELAKEVLSLEQIVLPCHPPEGI 176	
DB	121 FGLEIYWCQVWSSAGTTKERRAVIRIARLKNFDQEPGLKEVPLDHEVLVLCQRPPEGV 180	
QY	177 PPAEVEWLRNEDLVDPSLDPNVYITREHSLVVRQARLADTANYTCVAKNIVARRRSAAA 236	
DB	181 PVAEVEWLRNEDVIDPTQDTNFTLLTIDHLLIQRQARLSDTANYTCVAKNIVAKRSTTAT 240	
QY	237 VIVYUNGWSTWTEHVSVCASCGRWQKRSCTNPAPINGAFCEGQNVHDRTVSSLLV 296	
DB	241 VIVYUNGWSSWAEKSPSCNCRGQWQKRTCTNPAPINGAFCEGQ-AFOKTACTTIC 299	
QY	297 SVDSGSWPSKWSACGLDCTHWRSECDPAPNGGECQGTDLDTNCTSDLCVHSAG 356	
DB	300 PVDGAWTWSKWSACSTCAHRSCECAHRSCECAHRSCECAHRSCECAHRSCECAHRSCE 359	
QY	357 PEDVALYVGL-IAVAVCLVLLVLLVLYCRKKEGLSDVADSS-ILTSGFQPVSIKPSK 414	
DB	360 SGDAALYAGLVVAVFVVAIVVAILMAVGVVYRNCRDFDITDSSAALTGGFHPVNFXTAR 419	
QY	415 ADNPHLL--TTQPDLSITT--TYQSLCPROQSPKFOITNGHLLSLPLGGGHTLHHSSP 471	
DB	420 PSNPQLLHSPVPPDLTASAGIRGVPVYALQDS-TDKIPMTNSPLDPLPSLKVKYVSSST 478	

QY	472 T-----SEAEFVSRSTQNY-----FRS-----LPRGTNSMTYGTG 503	
DB	479 TSGGPGIADGADLLGLVLPPTPSDFARDTHFLHLSASIGSQQLLGLPRDPCGSSVSGTG 538	
QY	504 NFLGGRLMIPNTGISLLIPDAIPRGKIYEIYLTJLHKPEDVRLPLA-GCOTLLSPIVSCG 562	
DB	539 GCLGGRLSIPGTGVSLLVPGALPQGFYEMILLINKAEST-LPLSEGTQTVLSPSVTCG 597	
QY	563 PPGVLTFRVILLAMDHCGEFSFSDSWSLRLKKKSCGSEWQDVLHLGEEAPSHLYYCOLEA 622	
DB	598 PTGLLLCRPVILTMPCAEVSARDWIFQLKTAHQHWE-EVVTLDDEILNTFCYQCLEP 656	
QY	623 SACYVTEQIGFALVGEALSVAAAKRLKLLFPAPVACTSLEYNIRVYCLHDTHDALKEV 692	
DB	657 RACHILLDQLTGYVFTGESYSRSYSAVKRLQLAVAFALCTSLSEYSLRVYCLEDTPVALKEV 716	
QY	683 VOLEKQGGQLQCEPRVLHFKDSYHNLRSLSHDVSLSLWKSLLYSYQEIPEVHTWNGTQ 742	
DB	717 LELERTLGGYLVVEPKPLMFKDSYHNLRSLSHDLPFAHWRSKLLAKYQIPIFYHWSGSQ 776	
QY	743 RYLHCTFTLERSVSTSDLACKLWVQVEGQGSFISFNITKTDFRFAELLALESEAG-- 800	
DB	777 KALHCTFTLERSLASTELTCKICVQVEGEGQIFQLHTTLA-ETPAGSLDTLCSAPGST 835	
QY	801 VPALVGPSAFKIPFLIROKIISLDPDPCRRGADWRTLAQKHLDSHLFFASKPSPTAMI 860	
DB	836 VTTQLGPYAFKIPLSIRQKICNSLDAPNSRGNDRMLAQKLSMDRYLYNFATKASPTGVI 895	
QY	861 LNLWEARHPFNGNLSOLAAVAGLQGPAGLFTVSEAC 899	
DB	896 LDLWEALQDDGDGDLNSLASALEEMKSEMLVAVATDGDG 934	
	RESULT 14	
	Q86SN3 PRELIMINARY; PRT; 945 AA.	
ID	Q86SN3	
AC	Q86SN3; (TREMBlrel. 24, Created)	
DT	01-JUN-2003 (TREMBlrel. 24, Last sequence update)	
DT	01-OCT-2003 (TREMBlrel. 25, Last annotation update)	
DE	P53-regulated receptor for death and life.	
GN	P53RDL1	
OS	Homo sapiens (Human).	
OC	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;	
OC	Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.	
OX	NCBI_TaxID=9606;	
RN	[1]	
RP	SEQUENCE FROM N.A.	
RX	MEDLINE=22533857; PubMed=12598906;	
RA	Tanikawa C., Matsuda K., Fukuda S., Nakamura Y., Arakawa H.;	
RT	"p53RDL1 regulates of p53-dependent apoptosis.";	
RL	Nat. Cell Biol. 5:216-223(2003).	
DR	EMBL; AB096256; BAC57998.1; -	
DR	GO; GO:0004872; F:receptor activity; IEA.	
DR	GO; GO:0007145; P:signal transduction; IEA.	
DR	InterPro; IPR000488; Death.	
DR	InterPro; IPR003599; Ig.	
DR	InterPro; IPR007110; Ig-like.	
DR	InterPro; IPR003598; Ig c2.	
DR	InterPro; IPR000884; TSP1.	
DR	InterPro; IPR000885; TSP 1.	
DR	InterPro; IPR000906; ZUS_1.	
DR	Pfam; PF00531; death; 1.	
DR	Pfam; PF00047; Ig; 1.	
DR	Pfam; PF00090; TSP 1; 2.	
DR	Pfam; PF00791; ZUS_1.	
DR	PRINTS; PR01705; TSP1REPEAT.	
DR	SMART; SM00005; DEATH; 1.	
DR	SMART; SM00409; Ig; 2.	
DR	SMART; SM00408; IGC2; 1.	
DR	SMART; SM00209; TSP1; 2.	
DR	SMART; SM00218; ZUS; 1.	
DR	PROSITE; PS50835; IG LIKE; 1.	


```
DR PROSITE, PS50092; TSP1; 2.
KW Receptor.
SQ SEQUENCE 945 AA; 103637 MW; 56064E33F323447 CRC64;

Query Match
Best Local Similarity 51.8%; Score 2479; DB 4; Length 945;
Matches 493; Conservative 150; Mismatches 249; Indels 60; Gaps 18;

QY 1 MAVRGLWPALGIVLAAM--LRSGAQO-SATVANPVPGANPDLLPHFLVEPDVIV 56
Db 1 MGARGARGALLALLLWDRPRLSQAQTDGSEVLPSDFSPAPAEPLPYFLQEPDAYIV 60
QY 57 KNPVLLVCKVAPATQIFKCKNGEVRQVDHVIERTSGSGSEPTMEVRINVSQVKEV 116
Db 61 KNPVELRCRAPATQIFKCKNGEVRQVDHVIERTSGSGSEPTMEVRINVSQVKEV 120
QY 117 FGLEEYWCQVWSSSGTTSKQAYIRIARLRNFEQELAKEVSLQGIIVLPCRPPEGI 176
Db 121 FGLEEYWCQVWSSAGTTKSRAYVRIAYLRKNPQEPGLKEVPLDHEVLLQCRPEGV 180
QY 177 PPAEVEWLNEDLVDSLDNVITREHSLVVRQARLADTANYTCVAKNIVARRRSASAA 236
Db 181 PVAEVEWLNEDVIDTQDTNELLTDHNLIRQARLSDTANYTCVAKNIVARRSTTAT 240
QY 237 VIVYNGGWSWTWESVCSACGRGWKRSCTNPAPLNGGAFCEGQNVHRTVSSLLV 296
Db 241 VIVYNGGWSWAESVCSACGRGWKRSCTNPAPLNGGAFCEGQNVHRTVSSLLV 299
QY 297 SVDGSPWSPKWSACGLDCTHWRSECDPAPRNGGEECGQTDLTRNCTSDLCVHS--- 353
Db 300 PVDGAMTWSKWSACSTECARHWSRECMAPPQNGGRDCSGTLLDSKNTDGLCMQKKT 359
QY 354 -----ASPEDVALVGL-IAVAVCLVLLVLLVYCRKEGLDSVDSS-IL 401
Db 360 LSDPNSHLEASG--DAALYAGLVVAIFVVAAILMAVGVVYVRNRCDFTDITDSSAL 417
QY 402 TSGFQVSVKPSKADNPILL--TIQDLSITT--TYQSLCPRODGPSPKQFNTNHLGSP 458
Db 418 TGGFHPVNFKTARPSNPOLLHPSVPPDLTASAGIYGPVALQDS-TDKIPMTNSPLDP 476
QY 459 LGGGRHTLHSSPT-----SEAEFVSRLSTQNY-----FRS----- 490
Db 477 LPSLKVKVYSSTTSGPGLADGADLLGVLPPTGYPDFARDTHFLHRSASLGSQQLLG 536
QY 491 LPRGTENMTYGTNFIQGRLLMINTGISILLIPDAIPRGKIYIYILTKHPEDVRPLA- 549
Db 537 LPRDPGSSVSGTFCGLGRLSIFGTGVSLLVFNCAIPQGFYEMYLINKAEST-LPLSE 595
QY 550 GCOTLLSPIVSCGPPGVLLTRPVILAMDHGCEPSDPSWLSRLKKQCEGSEWQDVLHGE 609
Db 596 GTQTVLSPSVTCGTGULLCRPVILTPHCAEVSARDWIFOLKTAHQHWE-EVVLDE 654
QY 610 EAPSHLYYQLEASACVFTFQGRFALVGEALSVAARLKLKLLLPAPVACTSLEYNIRV 669
Db 655 ETLNTPCYQLEPRACHILLDQGTVFTGESYSRGAVKRLQLAVEAPALCTSLYSLRV 714
QY 670 YCLHDTHALKEVWQLEKQGLGQIIOBPVRLHFKDSYHNLRILSIHDVPSSLWSKLLVSY 729
Db 715 YCLEDTFVALKEVLELERTLGGVLEEPKPMFKDSYHNLRILSLHDLPHAHWSKLLAKY 774
QY 730 QEIPFHYHNGTORYLHCTFTLRSVPSSTDLACKLWQVQEGSGSFINFNITKDTRF 789
Db 775 QEIPFHYHNGSOKALHCTFTLRSVPSSTDLACKLWQVQEGSGSFINFNITKDTRF 833
QY 790 AELLALSEAG--VPALVGSAPKIPPLIRQKIISLDPCCRGADWRTIAQKLHDSHL 847
Db 834 GSLDTLCSAPGSTVTTQGLPYAFKIPLSIRQKICNSLDAENSGNDRMRLAQLKSLMDRYL 893
QY 848 SPFASKPSPAMILNLWEARHFNGLSQAALAVAGIQDAGLFTVSEAC 899
Db 894 NYFATKASPTGILLDLEALQDDGDLNSLASALEENGKSEMLVAVATDGC 945
```

```
RESULT 15
Q8K1S2
ID Q8K1S2 PRELIMINARY; PRT; 956 AA.
AC Q8K1S2;
DT 01-OCT-2002 (TREMELrel. 22, Created)
DT 01-OCT-2002 (TREMELrel. 22, Last sequence update)
DT 01-OCT-2003 (TREMELrel. 25, Last annotation update)
DE Netrin receptor Unc5h4.
GN UNC5H4.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
RN [1]
RP SEQUENCE FROM N.A.
RA Engelkamp D.;
RT "Cloning of three mouse unc-5 genes and their expression patterns at
RT mid-gestation.";
RL Submitted (MAY-2002) to the EMBL/Genbank/DBJ databases.
DR EMBL; AJ487854; CAD3252.1; -.
DR MGD; MGI:2389364; Unc5h4.
DR GO; GO:0004872; F:receptor activity; IEA.
DR GO; GO:0004867; F:serine protease inhibitor activity; IEA.
DR GO; GO:0007165; P:signal transduction; IEA.
DR InterPro; IPR000488; Death.
DR InterPro; IPR003599; IG.
DR InterPro; IPR007110; IG-like.
DR InterPro; IPR003598; IG_c2.
DR InterPro; IPR000215; Serpin.
DR InterPro; IPR000884; TSP1.
DR InterPro; IPR008085; TSP 1.
DR InterPro; IPR000906; ZUS.
DR Pfam; PF00531; death; 1.
DR Pfam; PF00047; ig; 1.
DR Pfam; PF00090; tsep_1; 2.
DR Pfam; PF00791; ZUS; 1.
DR PRINTS; PR01705; TSP1REPEAT.
DR SMART; SM00005; DEATH; 1.
DR SMART; SM00409; IG; 1.
DR SMART; SM00408; IGC2; 1.
DR SMART; SM00209; TSP1; 2.
DR SMART; SM00218; ZUS; 1.
DR PROSITE; PS50835; IG LIKE; 1.
DR PROSITE; PS00284; SERPIN; 1.
DR PROSITE; PS50092; TSP1; 2.
DR Immunoglobulin domain; Receptor.
SQ SEQUENCE 956 AA; 106351 MW; DDF07839C10C68D CRC64;

Query Match 44.8%; Score 2144.5; DB 11; Length 956;
Best Local Similarity 44.8%; Pred.No. 5.9e-189;
Matches 425; Conservative 162; Mismatches 283; Indels 79; Gaps 16;

QY 8 WPALGIVLAAMLRGS-----GACQSATVANPVPGANPDLLPHFLVEPDVIVKKEPVL 63
Db 15 WLPWLGLFF--WAAGAAAARGAGSGEILPDSIPSA-PGTLPHIEEPEDAYIIKSNPIAL 71
QY 64 VCKAVPATQIFKCKNGEVRQVDHVIERTSGSGSEPTMEVRINVSQVKEVKGLEEV 123
Db 72 RCKARFAMQIFKCKNGEVRQVDHVIERTSGSGSEPTMEVRINVSQVKEVKGLEEV 131
QY 124 CQCVWSSSGTTSKQAYIRIARLRNFEQELAKEVSLQGIIVLPCRPPEGIPPAEVEW 183
Db 132 CQCVWSSSGTTSKQAYIRIARLRNFEQELAKEVSLQGIIVLPCRPPEGIPPAEVEW 191
QY 184 LRNEDLVDSLDNVITREHSLVVRQARLADTANYTCVAKNIVARRRSASAAVIVYNG 243
Db 192 LKNEEPIDEQDNDIDTRADHNLIRQARLSDSGNYTCMAANIVAKRSLUSATVVIVYNG 251
QY 244 GNSWTWESVCSACGRGWKRSCTNPAPLNGGAFCEGQNVHRTVSSLLVSDGWS 303
Db 252 GNSWTWESVCSACGRGWKRSCTNPAPLNGGAFCEGQNVHRTVSSLLVSDGWS 310
QY 304 PWSKWSACGLDCTHWRSECDPAPRNGGEECGQTDLTRNCTSDLCV----- 351
```

Db 311 VMSWVCSPECHLRIRECTAPPNGGKFCGLSQESENCTDGLCILDKPLHEIKPQ 370
QY 352 -----HSASGPEDVALYGLIAVAVCLVLLVLLVILYCRKEGLDSDVADSSILTSGFPQ 407
Db 371 RNSRRGIEINASDIALYSGL-GAAVAVAVLIVIGITLYRRSHSDYGVVIDSSALTGGFQT 429
QY 408 VSIKPSKADNPHLL--TIQPDLSITTTTQGSCLCPRODGPSPKFLQTNHGLSLPLG----- 460
Db 430 ENFKTVRQNSULLNPFAMQPDLTVSTYSGPIC-LQD-PLDKELMTESSEFNPLSDIKVK 487
QY 461 -----GGRH-----TLHSSPTSEAEFVSRSLSTQNYER 489
Db 488 VOSSFMVSLGVSRAEYHGKXHSRGTPPHGNNRGFTIHPNKI-----PYIQNLS----- 537
QY 490 SLPRGTSNMTYCTFNLGGLMIPNTGISILLPPDAIPRGKIYEIYLTLHKPEDVRLPLA 549
Db 538 SLPTRELTGTGVFGHGLRLVMPNTGVSVLLIPHGAIPENSWEIYMSINQGEPS-SLOSD 596
QY 550 GCQTLLSPIVSGPPGVLLTRPVILAMDHCGBSPDSWSLRLLKQSCGSGWEQDVLHLGE 609
Db 597 GSEVLLSPEVTCGPPDMLVTTFALTIPHCADVSSSEHNIHLKRTQQQKWE-EVMSVED 655
QY 610 EAPSHLYYCOLBASACYVTEQIGRALVGEALSVAALKRLKLLPAPVACTSLEYNIRV 669
Db 656 ESTS--CYCLDPPFACHVLLDSFGYVALTGEPTDCAVKQLKAVVFGCMSCNSLDYNLRV 713
QY 670 YCLHPTHDALEKVVQLEKQGLQIQEPRVLHFKDSYHNRLRSIHDVPSLWKSLLVSY 729
Db 714 YCVDNTPCAFQEVISDERHQGQLUSEPKLLHFKGNTFSLQVSVLDDIPPFLWRIKPTAC 773
QY 730 QBIPTYHIWNGTQRYLHCTFTLERSVSPSTDACKLWVQVGDGQSFSGINFINITKDTRF 789
Db 774 QEVPSRVWSSNRQPLHCAFSLERYTPTTQJSCIKICIRQLKGHEQLQVQTSILEERE 833
QY 790 AELLALESEAGVPALVGPSAFKIPFLIRQKIISSLDPPCCRRGADWRTLAQKLHLDHLSF 849
Db 834 TITPFAQEDSTPFAQTGPKAFKIPYSIRQICATFDTNAGKGMQMLAQKNSINENLSY 893
QY 850 FASKPSPTAMILNWEARHFPNGNLSQLAAAVAGLQDPDAGLFTVSEAE 898
Db 894 FATQSSPSAVILNWEARHQDGLDLSLACALEEIRGTHTKLSNITEPQ 942

Search completed: October 4, 2004, 18:31:49
Job time : 148 secs

This Page Blank (uspto)